Totomac

il sistema TOTOMAC

LA NUOVA FRONTIERA DEL TOTOCALCIO PER C64

HOME E PERSONAL COMPUTER

GRUPPO EDITORIALE JACKSON

il sistema TOTOMAC

LA NUOVA FRONTIERA DEL TOTOCALCIO PER C64

Totomac

GRUPPO
EDITORIALE
JACKSON
Via Rosellini, 12
20124 Milano

INDICE

PREFAZIONE

Accattivarsi e dominare la fortuna è da sempre il sogno di ogni uomo, anche non giocatore d'azzardo. Riuscire a prevedere il futuro è l'altra grande aspirazione umana.

Il gioco del Totocalcio stimola a dimostrare tali capacità e Totomac affronta entrambi i compiti con una modalità originale: non inventa magie od esorcismi per accattivarsi la dea bendata, non è interessato a questa donna senza volto. Egli guarda in faccia e sviscera il comportamento dei rivali che gli contendono il bottino, e si inserisce con astuzia là dove "sistemisti" o "trottolisti" peccano di faciloneria e di ingenuità.

Uomo del nostro tempo, Totomac si affida all'ausilio del computer per individuare le probabilità meno utilizzate dalla massa dei giocatori, e, avendo ben presente i principi della selezione naturale, cerca di batterli con tecniche intelligenti.

Ma se, in futuro, tutti useranno i suoi metodi, cos'altro bisognerà inventare per battere la massa? Totomac non ce lo dice. Forse sta già pensando a nuove tecniche per battere anche se stesso...

E se non ci penserà lui, ci penserà qualcun'altro. Perché la voglia di vincere non finisce qui......

Ormai avevamo finito di cenare.

"Son tre mesi che ho incominciato a giocare al Totocalcio" – mi disse Baccarat – "E sto studiando un sistema per vincere".

Posò il tovagliolo e sorrise: "Potresti darmi una mano con il calcolatore, diventerebbe un sistema perfetto".

Baccarat è un tipo sempre pieno di iniziative, le novità della tecnica lo attraggono come calamite; che siano automobili State-of-the-art o Aziende Vinicole dell'Oltrepò, lui ne è al corrente. Si tratta di un ometto estroso, come potete vedere: lavora sul fronte delle comunicazioni, mass-media e via dicendo. Ama affrontare i problemi di cui va a caccia con un minimo di conoscenze di base. È per questo che mucchi di manuali minacciano di far crollare gli scaffali del suo studio.

Naturale quindi che volesse computerizzarsi per perfezionare il suo sistema di gioco: "Oggi come oggi – disse – non puoi sviluppare un nuovo progetto senza calcolatore".

Personalmente non sono un giocatore, credo fermamente nella eventuale cecità della fortuna e non mi diverte calcolare le possibili combinazioni di serie di numeri. E poi, perbacco, mettermi a fare giochetti da home-computer! Sul nostro, tutto sommato, serio e professionalissimo calcolatore! Non era un'idea stimolante.

Però il Totocalcio è un argomento fecondo di temi interessanti = soldi, calciatori e, se vuoi mettere a punto un sistema, statistica: mi pareva che l'uditorio fosse particolarmente reattivo quella sera.

Eravamo in cinque. Di fronte a Baccarat sedeva un grosso matematico olandese, Van Random, che insegna statistica, di fronte a me sedeva Totomac: un fisico che lavora, stranamente, nell'ufficio studi di una finanziaria. Il quinto dormiva sotto il tavolo ed aveva le sembianze del bassotto di Baccarat.

Mentre ci trasferivamo sulle poltrone tentai comunque di scansare il problema, dichiarandomi completamente ignorante sui diversi sistemi di giocata. Questo non era un ostacolo per Baccarat: "Ah, ti faccio vedere subito – e tirò giù da uno scaffale una pigna di giornali e di libri sulle cui copertine brillavano invitanti biglietti da centomila – ci sono mol-

tissimi sistemi, perché bisogna tenere conto di moltissime cose per poter vincere.

Per esempio devi calcolare la probabilità di avere più o meno 1, X e 2 in ogni colonna, poi può darsi che un risultato non esca da diverse settimane e allora devi aspettarti che salti fuori da un momento all'altro, senza contare le doppie e le triple". E andò avanti per dieci minuti buoni parlando di formule derivate e facendo spesso riferimento alla teoria dei grandi numeri.

Van Random, intanto, sfogliava i manuali posati sul tavolino. È un tipo sostanzialmente pestifero, benché intelligente e a volte simpatico. Pensa che i numeri siano una sua personale proprietà e considera chi osa avvicinarsi ad essi senza un'adeguata preparazione come un malintenzionato grossolano e pericoloso. È pronto ad aggredire chi abbia il coraggio di farlo, quella sera esplose con particolare violenza.

"Tutto questo è intollerabilmente ridicolo – e il suo tono era definitivo – mi meraviglio che, dopo tanti anni che hai me per amico, tu faccia ancora questi discorsi. Quelle che stai dicendo son storie che vanno bene per le vecchiette che giocano al lotto. Se c'è una cosa seria nel Totocalcio, è che, bene o male, si sa se una squadra è più forte di un'altra. L'unica cosa da fare è valutare questa differenza e cercare di indovinare chi vincerà. Il resto son tutte balle".

"Ma io seguo molto le squadre..." – diceva Baccarat.

"E allora cosa c'entrano queste altre storie – lo interruppe Van Random senza la minima discrezione –, tutti sanno che 'la pallina non ha nè memoria nè coscienza'. Questo vale per la pallina della roulette e altrettanto per il pallone da calcio. Cosa c'entrano i 'ritardi'. Due squadre in campo giocano come giocano; cosa c'entrano i totali di 2 o di X in una schedina. Certamente le vittorie in trasferta sono più difficili e quindi in genere i 2 sono pochi, ma se un giorno quattro o cinque squadre possono e si trovano a vincere fuori casa certamente non si mettono a perdere perché nella schedina non escano troppi 2. Discorsi del genere sveglierebbero Bernoulli".

"Però il Totocalcio mi sembra diverso" – azzardò improvvisamente Totomac ma Baccarat era stato toccato troppo sul vivo e lo interruppe per replicare a Van Random.

D'altronde Totomac non possiede l'arte della conversazione e non finisce quasi mai i discorsi che incomincia.

I suoi processi mentali sono di tipo markoviano con barriere assorbenti sparse qua e là e solo a tratti i suoi circuiti elaborano un risultato

comunicabile. Questo affiora in output come una testuggine sulla superficie del mare, rimane inerte per un po' poi si riimmerge per essere riprocessato. Tutt'altro tipo Baccarat, i cui pensieri continuano a saltare dentro e fuori come delfini per ossigenarsi, e che riprese: "Ma lo sanno tutti che dopo una serie di rossi deve uscire il nero. Allo stesso modo, dopo una serie di 1 deve uscire un risultato diverso. Perché alla lunga la differenza fra i risultati tende a zero. È la teoria dei grandi numeri".

"Ma per carità – sbottò Van Random, congestionato e furente – non è affato vero. Quella che tu chiami la teoria dei grandi numeri non si è mai sognata di dire una cosa del genere. Questi sono discorsi da mezzopollo-romanesco".

Spesso l'ho sentito esprimersi con durezza sui mezzi-polli-romaneschi, che sono per lui i prototipi dell'ignoranza in materia di statistica. "Comunque – continuò un po' più disteso – è vero che il risultato converge alla media, o se vuoi che lo scostamento dalla media tende a zero, ma solo perchè la media è data dalla somma dei risultati diviso il numero degli esperimenti che si considera tendente all'infinito. Ma la differenza assoluta tra il numero dei rossi e dei neri non va affatto a zero. Anzi, se mai, tende ad aumentare. È molto più probabile, per esempio, avere una differenza di 10 a favore dei rossi dopo 1000 lanci che non dopo soltanto 100. Questo è quel che dice la teoria e nessun esperimento al mondo lo ha mai smentito".

"Non ci può essere nessun sistema per vincere – concluse Van Random – salvo quello di pronosticare bene e azzeccare il risultato. Tutti i sistemi che ho visto nei tuoi libri servono o a compilare più rapidamente le schedine, oppure a spostarti verso valori di rischio più bassi diminuendo evidentemente anche il valore atteso della vincita, il che è comunque un obiettivo più che legittimo, oppure sono superstizioni da vecchiette che, come ti dicevo, farebbero fare a Bernoulli un salto altro tre metri.

Nessuno migliora realmente l'aspettativa del tuo gioco".

Baccarat era temporaneamente bloccato e Totomac approfittò dell'attimo di impietrito silenzio prodotto dalla franchezza di Van Random.

"Mi sembra che il problema sia diverso – disse – perché il Totocalcio non è la roulette e d'altra parte, 'pronosticare bene' non è mica facile. Però ci può essere un sistema: è quello di giocare la gamma più remunerativa".

"Cos'è la gamma più remunerativa" – chiese Baccarat interessato alla parola.

Totomac era evidentemente riuscito a dare una forma comunicabile alle sue riflessioni. "La sostanza del discorso – rispose – è che al Totocalcio la vincita non è fissa come alla roulette o al lotto ma risulta dal numero di giocatori che hanno pronosticato esattamente e che si dividono le puntate dei giocatori che hanno sbagliato. Quindi si tratta di puntare sui risultati che gli altri giocatori rifiutano; però non in modo cieco perché il risultato al calcio ha una sua logica e certi risultati non usciranno mai, ma facendo i conti preventivamente".

"Supponiamo per esempio – continuò – che andiate ai Caraibi. Lì fanno i combattimenti dei galli. Supponiamo che esista una specie di Totocalcio ma molto più semplice: si fa il pronostico su un solo incontro anzichè su tredici ed esistono solo vittoria e sconfitta, non c'è pareggio. Il monte premi, e questo è il primo fatto importante, viene diviso come al Totocalcio fra quelli che hanno pronosticato il risultato giusto. Inoltre, e questo è il secondo fatto importante, come al Totocalcio tutti in generale sono d'accordo su quale dei due galli sia più forte e tutti sanno anche che effettivamente il favorito vince quasi sempre. Quindi puntando sul favorito si vince spesso ma poco, mentre puntando sulla vittoria dell'altro si vincerà più di rado ma il premio sarà più alto. Benissimo, fin qui tutto è ovvio, però ora si tratta di fare i conti".

"Prendiamo i risultati degli ultimi concorsi, diciamo dell'ultimo anno – continuò – di cui possiamo avere monte premi, e quindi il numero dei giocatori, ed il numero dei vincitori. Ora immaginiamo di trovare che nel 60% dei casi il premio è stato diviso fra l'80% degli scommettitori e che nel 40% dei casi il premio è stato diviso fra il 20%. Bene quello che avremmo trovato non disturba affatto il sonno di Bernoulli ma riguarda soltanto il comportamento degli scommettitori: mentre la probabilità di vittoria del presunto favorito è solo del 60% questo viene puntato dall'80% dei giocatori. Questo 'sovrapuntare' il favorito può essere un fatto psicologico, economico, o di propensione al rischio; in ogni caso si tratta di un comportamento sistematico e costante dei giocatori che però è viziato e sfruttando questa strategia viziata tu puoi trovare la tua strategia vincente. Infatti è immediato, tu devi puntare il risultato 'sottopuntato' dagli altri, cioè contro il favorito.

Supponiamo che ci siano normalmente 100 scommettitori con una puntata di 100 lire e che il banco non prelevi nulla dando quindi un monte premi di 10.000 lire.

Dopo 100 settimane avrai puntato complessivamente 10.000 lire e avrai vinto 40 volte dividendo la vincita con altre 20 persone, incassando quindi 40×10.000/20=20.000 lire e sarai in vincita; se invece avessi puntato il favorito avresti vinto 60 volte ma dividendo la vincita con altre 80 persone con un risultato di 60×10.000/80=7.500 lire e saresti quindi in perdita".

"Va da sè – continuò Totomac – che in un caso così semplice anche gli scommettitori dei Caraibi se ne sarebbero accorti e avrebbero adottato strategie più equilibrate. Il Totocalcio però è più complesso e non è facile rilevare una eventuale situazione di questo genere. In ogni modo, se da un esame dei risultati passati dovessimo trovare che anche nel Totocalcio esistono sistematicamente colonne 'sovrapuntate' e colonne 'sottopuntate', questo potrebbe costituire la base di un 'sistema'".

"Mica scemo – commentò Van Random con la consueta amabilità, ma Baccarat non era affatto convinto". "Un momento – obiettò – tu dici che si deve puntare il risultato 'eccezionale', ma questo è quello che fanno in molti, ma non vincono mai".

Personalmente, ero d'accordo con lui.

"No, non è questo che dice Totomac – intervenne Van Random la cui velocità di elaborazione è portentosa e che si riteneva evidentemente più qualificato di Totomac a sostenere la discussione –, non è detto che la 'gamma più remunerativa', come dice lui, sia quella dei risultati meno probabili, anzi verosimilmente non lo è. Può essere quella dei risultati più probabili, oppure una gamma intermedia. Comunque è quella dei risultati la cui probabilità è maggiore della percentuale di puntate che ricevono".

Van Random è un leader naturale: all'inizio critica tutto, ma appena la cosa mostra di poter marciare, arraffa.

"Perché – proseguì accalorandosi man mano che parlava – la sostanza è che in questo gioco l'avversario non è il caso, contro il quale non hai speranza, ma sono gli altri scommettitori; è lo scommettitore, inteso come la somma di tutti quelli che puntano al Totocalcio".

"E questo scommettitore – Van Random era sempre più lanciato – una strategia ce l'ha, e se tu la trovi puoi giocarci contro con successo. Anzi, meglio ancora, ce l'ha ma non la può controllare, perché lo scommettitore è fatto da migliaia, anzi milioni, di teste che ragionano ognuna per conto proprio. Tra l'altro, questa strategia ha l'aria di essere piuttosto stabile, perché milioni di giocatori non cambiano abitudini da un giorno all'altro".

"Scusa ma non ho capito bene – intervenne ancora Baccarat che aveva ascoltato grattando distrattamente la testa del bassotto addormentato. "Ma è chiarissimo – Van Random respinse l'obiezione con noncuranza – ai Caraibi, la 'gamma' del favorito succhia ogni settimana 8000 lire e ne paga mediamente 6000, mentre la 'gamma' dell'outsider ne succhia 2000 e ne paga 4000".

"Ah interessante – acconsentì Baccarat ancora dubbioso –". "Oppure, lasciando stare i Caraibi e tornando alla roulette – proseguì Van Random che con la solita perspicacia aveva trovato il canale giusto per comunicare con Baccarat – immaginati una roulette in cui valga la regola di dividere il monte delle puntate fra i vincitori anziché avere un premio fisso. Supponi che siano usciti 100 rossi di fila e che tutti i giocatori che aspettano il ritardatario hanno coperto il nero di fiches e il rosso è completamente vuoto. Tu cosa punti?".

"Beh – rispose Baccarat visualizzando la scena con un certo entusiasmo – È ovvio, punto il rosso".

Eravamo ormai lanciati e fino a notte avanzata discutemmo i vari aspetti del sistema, finché non mi azzardai a proporre: "Bene, visto che non abbiamo svegliato Bernoulli, perché non andiamo a dormie anche noi". Il seguito è nei prossimi capitoli.

Salutandoci sulla porta, Baccarat si chiese perché chiamavamo Bernoulli il suo cane.

INTRODUZIONE

IL SISTEMA TOTOMAC IN SINTESI

Il sistema Totomac è composto di tre sezioni:

— la prima sezione riguarda il modello ed i programmi di calcolatore necessari a stimare il numero di puntate che riceverà una determinata qualsiasi colonna (frequenza di giocate).
  Questo serve a determinare con quanti altri vincitori dovreste dividere la vincita ed è trattato nella prima parte del libro, nella quale è riportato il modello desunto dai concorsi della stagione 83-84 e sono riportati i criteri ed i programmi per aggiornare il modello stesso.

L'ultimo capitolo della parte terza riprende alcune indicazioni per il miglioramento ed aggiornamento del modello stesso.

— La seconda sezione riguarda il calcolo della probabilità di uscita di una determinata qualsiasi colonna ed è trattato nella parte seconda.
  Questo richiede un'assegnazione preventiva di probabilità a ciascun risultato di ciascuna partita. Se vi ritenete un buon pronosticatore questa assegnazione di probabilità potete farla voi stessi, se non vi ritenete tale, potete prendere un qualsiasi pronostico in forma standard pubblicato da un quotidiano o rivista specializzata.
  Nella parte seconda sono contenuti programmi per la taratura e la valutazione del pronosticatore e per lo sviluppo del pronostico. Naturalmente questi programmi ignorano se il centravanti si è fratturato e quindi servono solo a chi non si senta in grado di fare di meglio da solo.

— La terza sezione riguarda la metodologia ed i programmi per produrre colonne che rispondano ai requisiti di essere colonne «buone», in base al rapporto fra probabilità e frequenza di giocate ed è trattata nella parte terza.
  I programmi consentono di definire la vostra strategia, secondo i vostri obiettivi e desideri, fissando il numero di colonne che volete giocare. I livelli di probabilità e di "remuneratività", la "riduzione" del sistema etc..

I programmi contenuti in questo volume sono scritti in Basic per il Commopdore 64; sono stati scritti originariamente per un mini Data General e successivamente trascritti per adattarli alle caratteristiche del C/64.

Il sistema Totomac richiede l'uso di un calcolatore. Benchè qualunque cosa faccia un calcolatore, come ormai è stato accertato, possa essere fatta anche da una mente umana, i tempi di generazione delle giocate sarebbero inaccettabili senza un mezzo automatico. Inoltre il sistema Totomac produce colonne indipendenti, senza passare attraverso sistemi completi, e quindi anche il controllo dei risultati riuscirebbe molto accelerato memorizzando le giocate e controllandole automaticamente, questa funzione non è inclusa nei programmai inseriti in questo volume.

PARTE I

IL MODELLO TOTOMAC

CAPITOLO 1

SULLA POSSIBILITÀ DI SVILUPPARE UN SISTEMA PER IL TOTOCALCIO

"Equità" di un gioco

Come abbiamo visto nella introduzione, un sistema per il Totocalcio è, in linea di principio, possibile in quanto il Totocalcio non è necessariamente un gioco "equo", come non lo sono tutti i giochi basati sul criterio del totalizzatore e cioè sulla ripartizione tra i vincitori delle quote giocate dai perdenti.

Non c'è, evidentemente, in questo nessun giudizio morale, che non essendo "equo" debba essere considerato "iniquo", si tratta soltanto delle regole del gioco stesso.

Definiamo infatti, per quanto ci riguarda, come gioco "equo" semplicemente un gioco nel quale le puntate che hanno uguale probabilità di vittoria ricevono un premio uguale, o più in generale in cui il premio è inversamente proporzionale alle probabilità di successo.

In questo senso è "equa" la roulette: qualunque numero, se esce, paga 36 volte la posta, o ugualmente è "equo" il lotto: qualunque terno, se esce, paga lo stesso premio.

Al Totocalcio, diversamente da questi, il premio non è fisso, ma dipende dal numero di giocatori che hanno puntato sul risultato uscito e pertanto è "equo" solo se lo rendono tale gli scommettitori stessi. Infatti, risulterà "equo" se gli scommettitori distribuiranno le loro puntate fra le diverse colonne in modo proporzionale alla loro probabilità di uscita e non lo sarà se si comporteranno diversamente, distribuendo le loro puntate in modo diverso.

Se il gioco risulta "equo", non abbiamo nessuna possibilità di sviluppare un sistema. Se invece il gioco non risulta "equo" e se riusciamo ad

individuare la strategia seguita dagli scommettitori nel distribuire le loro puntate, possiamo sviluppare un sistema vincente.

Il problema di Totomac è quindi, prima di tutto, verificare se il gioco sia "equo" e successivamente, se non lo è, individuare la strategia degli scommettitori. Cioè individuare quel tipo di colonne che saranno puntate dagli scommettitori in misura inferiore alla loro probabilità di uscita. In questo modo, Totomac si propone di migliorare quello che viene chiamato il "valore atteso" di una puntata, che viene espresso da:

valore atteso = probabilità di vincita * valore della vincita – costo

Con un esempio numerico, supponendo di puntare 100 lire a testa e croce e di incassarne 200 in caso di vincita, il valore atteso è:

VA = 0,5 * 200 – 100 = 0

Come chiunque avrebbe potuto già anticipare senza troppi calcoli, giocando a testa e croce, mediamente c'è da aspettarsi di chiudere in pari.

Al Totocalcio, essendo la vincita uguale al monte premi diviso il numero dei vincitori, il valore atteso sarà:

VA = probabilità della colonna * monte premi / numero delle puntate sulla colonna – costo

In conseguenza, se la probabilità della colonna ed il numero di puntate su di essa stanno sempre nella stessa proporzione, il gioco è "equo" ed il valore atteso è uguale per tutte le colonne, se invece possiamo individuare dei "buchi", cioè delle colonne per cui la frequenza di giocata è più bassa rispetto alla probabilità, per queste colonne avremo un valore atteso più alto.

L'obiettivo di Totomac è appunto di migliorare il valore atteso del giocatore, diversamente dagli altri sistemi che non si propongono direttamente questo risultato.

Verifica della "equità" del Totocalcio

Per verificare se il Totocalcio sia o no "equo", dobbiamo quindi confrontare le probabilità di uscita delle colonne con il numero di puntate che esse effettivamente raccolgono.

Per questo possiamo rifarci ai concorsi passati: le informazioni di cui disponiamo non sono molte, in pratica possiamo prendere in esame le sole colonne vincenti e mettere a confronto:

- numero di giocate raccolte dalla colonna vincente
- probabilità della colonna stessa.

PARTITA N.	1	X	2
1	40	40	20
2	25	35	40
3	60	30	10
4	33	33	33
5	75	20	5
6	45	35	20
7	30	50	20
8	50	30	20
9	60	30	10
10	45	30	25
11	50	35	15
12	33	33	33
13	40	30	30

Figura 1

Per valutare quale probabilità poteva essere attribuita alla colonna vincente dobbiamo rifarci ad un pronostico, formulato evidentemente prima della partita ed espresso in forma probabilistica standard, e cioè del tipo riportato in figura 1, che attribuisce un valore percentuale di probabilità a ciascuno dei tre risultati.

La probabilità di ciascuna colonna è data dal prodotto delle probabilità dei tredici risultati verificatisi. Ad esempio, per il pronostico della Figura 1, la colonna 1111111111111 avrebbe probabilità:

,40 × ,25 × ,60 × ,33 × ,75 × ,45 × ,30 × ,50 × ,60 × ,45 × ,50 × ,33 × ,40 = 0.000017862

e cioè 17.8 per milione.

Se questa colonna viene giocata da 18 giocatori per milione è una colonna indifferente, se viene giocata da più di 18 giocatori per milione è una colonna da evitare, se viene giocata da meno di 18 giocatori per milione è una colonna buona.

Tanto per fare l'occhio alle cifre in gioco, 18 giocatori per milione significano, in una settimana normale da 150 milioni di colonne giocate, 2700 eventuali vincitori su quella colonna.

Le informazioni utilizzate

Sono stati presi in esame, per la verifica e lo sviluppo del modello, 21 concorsi dell'84, e precisamente i concorsi numero: 6,7,10,11,12,15,16,17,18,20,21,22,23,25,26,27,29,30,31,32,34 ed i relativi pronostici pubblicati dalla "Repubblica" il sabato precedente le partite.

L'intero set dei dati: numero di giocate, numero dei tredici, colonna vincente e pronostico è riportato nella Tabella 1.

La scelta di questi concorsi è stata fatta dai dati di vincita pubblicati dal Totoguida del 24/4/84. Sono stati esclusi i concorsi numero 4 e 5 in quanto mancava il pronostico sulla "Repubblica", tutti i concorsi di se-

rie B e Coppa Italia ed il concorso numero 19 nel quale non si sono verificati tredici.

Purtroppo hanno dovuto essere esclusi anche i concorsi numero 8 e 14. Respingiamo con fermezza l'eventuale insinuazione che questi siano stati esclusi "a posteriori" per far tornare a forza i risultati del modello. La realtà è che, al momento di abbinare il pronostico ricavato dalla "Repubblica" al risultato ricavato dal "Totoguida", sono sorti dubbi se il pronostico trascritto fosse relativo al concorso voluto oppure ai concorsi adiacenti numero 9 e 13 di Serie B.

Totomac, responsabile di non avere trascritto i riferimenti necessari per effettuare senza dubbi l'abbinamento fra pronostico e risultati, si è comunque rifiutato di passare un'altra mezza giornata alla biblioteca comunale di Parco Sempione a sfogliare numeri arretrati della "Repubblica". Ci disse che ricordava come un incubo il sabato mattina che ci aveva già passato, sommerso da liceali, con gravi difficoltà per rifare i pacchi dei giornali dopo averli sfogliati e per riannodare gli spaghi forniti dal comune. Ci disse che quegli spaghi gli ricordavano sua nonna che in vita sua non aveva mai tagliato lo spago di un pacco, ma disfaceva accuratamente tutti i nodi per conservare lo spago per casi di necessità. Ci disse che la direttrice era stata gentilissima, ma che le aveva visto nello sguardo il sospetto di un basso fine occulto, come la partecipazione ad una caccia al tesoro o qualcosa del genere e che in definitiva non sarebbe tornato a chiedere ad una povera ragazza di tirare fuori altre tonnellate di carta dai sotterranei.

Abbiamo concluso che ventuno concorsi ci sarebbero bastati.

La verifica delle ipotesi

Nel diagramma di Figura 2 sono riportati i dati dei 21 concorsi considerati in una forma opportuna per dare una prima occhiata.

Nel nostro progetto di empirismo-assistito-da-calcolatore, è comunque un accorto primo passo, quello di diagrammare i dati e considerare a vista come si presentano.

I dati sono riportati in scala bilogaritmica; non c'è un particolare motivo per fare questo, se non il fatto che riportare contemporaneamente nello stesso grafico valori dell'ordine di grandezza delle diecine di unità

e delle centinaia di migliaia è impossibile se non si usa la scala logaritmica. Il grafico logaritmico va però letto in un modo particolare, che vedremo più avanti.

Per ciascun concorso, è stata calcolata la probabilità assegnata dal pronostico della "Repubblica", alla colonna effettivamente vincente e questa, moltiplicata per il numero totale delle colonne giocate in quel concorso fornisce il numero dei vincitori "equi", cioè il numero di vincitori che avremmo avuto se le puntate fossero state distribuite fra le colonne in base alla loro probabilità relativa.

L'asse delle ascisse rappresenta appunto il numero di vincitori "equi", mentre l'asse delle ordinate rappresenta il numero dei vincitori effettivi. Ogni concorso è rappresentato da un punto nel diagramma.

La retta centrale a 45 gradi rappresenta la retta di indifferenza: su questa retta si troverebbero i concorsi in cui il numero di vincitori "equi" fosse uguale al numero dei vincitori effettivi; tutti i punti al di sopra di questa retta rappresentano concorsi in cui la colonna vincente è stata "sovrapuntata", cioè ha ricevuto più puntate di quante comportasse la sua probabilità di uscita, al contrario, i punti al di sotto della retta rappresentano concorsi in cui la colonna vincente è stata "sottopuntata".

La prima occhiata al diagramma è decisamente incoraggiante dal nostro punto di vista: circa metà delle colonne vincenti risultano "sovrapuntate" e circa metà risultano "sottopuntate" e gli scostamenti rispetto al valore di indifferenza sembrano rilevanti. Nel diagramma logaritmico le rette parallele rappresentano punti che stanno fra loro in un rapporto costante: la retta contrassegnata con 10, ad esempio, rappresenta valori di puntate effettive dieci volte superiori ai valori equi, analogamente la retta contrassegnata con 1/4 rappresenta valori di puntate effettive pari ad un quarto dei valori equi. Saremmo tentati di concludere subito che effettivamente nel totocalcio la distribuzione delle puntate non segue la probabilità delle puntate stesse.

Tuttavia, ci possono essere delle obiezioni:

La prima è che il pronosticatore della "Repubblica" sia un mediocre pronosticatore o che comunque non dia la stessa valutazione di probabilità data dalla massa degli scommettitori. In effetti nessuno ha mai visto copie della "Repubblica" appese nelle ricevitorie, gli scommettitori sembrano avere altre letture.

Inoltre la stessa persona che ha formulato il pronostico, se messa alle strette, avrebbe difficoltà a sostenere che la probabilità di un risultato fosse proprio del 40% e non piuttosto del 35% o 45%.

La risposta a questa obiezione può essere che dopo tutto si tratta di uno staff di giornalisti rispettabili e che, per l'altro aspetto, le incertezze nella probabilità dovrebbero tradursi in scostamenti più forti nei concorsi con colonne vincenti a bassa probabilità, che contengono parecchi risultati a bassa probabilità, e più limitate in quelli a probabilità alta, il che però dal diagramma non risulta.

Infatti, è credibile che un pronosticatore, anche accurato, assegni ad un risultato la probabilità del 5% mentre la massa degli scommettitori, o la realtà oggettiva (qualunque cosa ciò significhi) gliene assegnano il 10%: questo solo fatto porterebbe a raddoppiare il valore globale di probabilità della colonna. È meno credibile invece che un pronosticatore coscienzioso sia in disaccordo con la massa assegnando una probabilità del 60% contro un 80%, che dopo tutto darebbe solo una differenza del 25% nel risultato finale.

La difesa finora è piuttosto debole, ma la risposta determinante è un'altra: 'ebbene, se il pronosticatore che abbiamo scelto è un pazzo inaffidabile che spara pronostici a casaccio, il nostro tentativo di trovare la strategia degli scommettitori basandoci su di lui semplicemente fallirà. In ogni modo non ne cercheremo altri'.

In questo modo siamo dalla parte della corretta metodologia: meglio non trovare una eventuale strategia esistente che riservarci troppi gradi di libertà nella ricerca e credere, alla fine, di aver trovato qualcosa mentre in realtà si sono soltanto tagliate le gambe dei dati su un letto di lunghezza troppo corta.

Una seconda obiezione è doverosa: si può sostenere che gli scommettitori giochino effettivamente secondo probabilità uguali o molto prossime a quelle della "Repubblica" ma che gli scostamenti fra equo ed effetivo siano dovuti soltanto al caso. Questa obiezione vanificherebbe qualunque ricerca di strategia.

In effetti, se la probabilità di una colonna è, ad esempio, di 1 su un milione con 150 milioni di colonne giocate potremmo aspettarci di trovare "mediamente" 150 giocate, ma non di trovare ogni volta esattamente e sistematicamente 150.

Per verificare questa ipotesi abbiamo però degli strumenti precisi e per nostra fortuna un campione così grosso come raramente è dato di avere.

Dobbiamo eseguire un curioso ragionamento a senso unico che va sotto il nome di "test dell'ipotesi nulla", che ha però tutta l'aria di essere l'unico fondamento della conoscenza, anche al di fuori del totocalcio.

Prendendo in esame un particolare concorso, ad esempio il numero 7, ci chiediamo:

> "Se la probabilità di compilare la colonna vincente (1211X21X211XX) da parte di uno scommettitore è quella 'equa' di 0.0000002052864, e questa operazione viene ripetuta 133.282.314 volte (quante sono le colonne giocate), qual'è la probabilità di ottenere per caso (questa è la ipotesi nulla) almeno 66 colonne di questo tipo (quante ne sono state effettivamente giocate)?"

Se la probabilità è "alta", diciamo almeno lo 0.5%, non ci sentiamo autorizzati a concludere che lo scostamento riscontrato sia dovuto ad altro che al caso, se invece la probabilità è "bassa" (inferiore al limite prescelto) possiamo ritenere "falsificata" l'ipotesi nulla e quindi respingerla e trarre la modesta conclusione che: "esistono componenti non-casuali che determinano lo scostamento dello scommettitore dallo schema della equità".

Potremmo quindi passare al capitolo 2 a cercare di individuare queste "componenti non-casuali" che determinano la strategia dello scommettitore.

Il ragionamento sembra a prima vista francamente demenziale.

Tuttavia dobbiamo tenere presente che i dati empirici possono dimostrare falsa una ipotesi, ma mai dimostrarla vera. Il massimo a cui possiamo arrivare è di dire che "non abbiamo motivi per ritenerla falsa".

Quindi ci dobbiamo cautelare assumendo come ipotesi "privilegiate" quelle ipotesi che, se false, possano essere dimostrate tali, ed evitando accuratamente ogni ipotesi che non abbia queste caratteristiche.

È come se, per quanto vale una analogia, per raggiungere il cancello di imbarco del nostro aereo dovessimo dimostrare non di aver diritto di andare dove vogliamo, ma piuttosto di non aver diritto di andare altrove.

Da Pangloss in qui, in ogni modo, lo sbarco di clandestini è stato tanto importante quanto l'imbarco di nuovi passeggeri.

I calcoli possono essere effettuati con il programma allegato.

Per i nostri dati, risulta che anche per concorsi come il 7, il 30 od il 31, che sono fra i più vicini alla retta di indifferenza, lo scostamento è tale che possiamo tranquillamente respingere l'ipotesi nulla ed accettare l'esistenza dei fattori non - casuali.

In definitiva, il totocalcio non è "equo", una strategia dello scommettitore esiste, non rimane che trovarla.

Per completare la illustrazione della figura 2, le spezzate tratteggiate rappresentano, grossolanamente, l'andamento dalle due parti della retta di indifferenza, delle curve di uguale "significatività" dello scostamento dalla condizione di indifferenza stessa: i punti situati su una di queste spezzate rappresenterebbero gli scostamenti ugualmente "significativi" al variare della probabilità.

CAPITOLO 2

L'IDENTIFICAZIONE DELLA STRATEGIA DELL'AVVERSARIO

Premessa

Questo capitolo riguarda lo sviluppo del metodo per la determinazione della strategia degli scommettitori, che si concreta nel numero di puntate per ciascun tipo di colonna.

Chi desideri sviluppare un proprio modello o comunque desideri approfondire la metodologia di sviluppo, può trovarlo interessante; d'altra parte il capitolo stesso può essere saltato.

Il modello definitivo è esposto al capitolo successivo, e quindi chi sia interessato soltanto alla applicazione del metodo può passare tranquillamente al capitolo 3.

Considerazioni generali

Nel capitolo precedente abbiamo accertato, con un buon margine di probabilità di non essere in errore, che la frequenza con cui una colonna viene giocata è sufficientemente diversa dalla probabilità di uscita della colonna stessa da escludere che queste differenze siano dovute al caso.

Esiste cioè un comportamento dello "scommettitore"; il nostro problema è ora di trovare, se ci riusciamo, la regola o le regole sistematiche che segue il nostro avversario.

Dobbiamo cioè trovare il "modello" che rappresenta, in funzione dei dati che conosciamo prima della domenica, la frequenza di puntate su

ciascuna delle 1.594.323 colonne possibili.

Il numero di puntate è la nostra variabile "obiettivo", i dati che sono disponibili prima della domenica sono delle possibili variabili "causa".

Un problema di questo genere viene affrontato con una procedura standard per successive fasi:

FASE 1

Scegliere le variabili che riteniamo importanti.

Questa scelta dipende esclusivamente dal nostro criterio e dalla nostra fantasia. Nessuna variabile può essere esclusa in linea di principio.

Chi può escludere, ad esempio, che se il tempo è piovoso gli scommettitori si comportino diversamente da quando il tempo è bello? Dobbiamo verificare sui nostri dati se queste variabili sono importanti o no.

Comunque, il buon senso ci aiuta senz'altro a scegliere delle variabili "causa" significative. È inutile elencare a casaccio variabili che "potrebbero" essere rilevanti, è molto meglio riflettere sul possibile modo di ragionare degli scommettitori e sul meccanismo con cui possono essere prodotte le puntate e definire le nostre variabili "causa" con una certa logica.

Saranno sempre abbastanza.

Prendiamo quindi la variabile che riteniamo più importante e passiamo ad esaminare se effettivamente lo è.

Esistono delle tecniche numeriche per verificare se esista una "correlazione" fra due variabili. Queste tecniche corrispondono grosso modo al procedimento grafico di riportare i punti relativi ai nostri dati in un diagramma in cui l'asse delle ordinate rappresenta la variabile "obbiettivo" da determinare e l'asse delle ascisse rappresenta la variabile "causa" e controllare se la nuvola dei punti è allungata ed inclinata rispetto agli assi, oppure no.

Se è praticamente circolare le due variabili sono "indipendenti", la variabile "causa" non è in realtà influente per determinare la variabile "obiettivo"; dovremo quindi scartarla e considerarne un'altra.

Se la nuvola dei punti è allungata come una ellisse, ma con gli assi paralleli ad uno degli assi coordinati, è chiaro che basta un cambiamento nella scala del disegno, per vederla diventare circolare.

Se invece la nuvola è allungata ed inclinata rispetto agli assi, possiamo concludere che le due variabili sono "correlate" fra loro, che la variabile "causa" è importante e significativa e possiamo passare alla fase 2.

Le tecniche numeriche cui abbiamo accennato, più o meno sofisticate, vanno sotto il nome di analisi di "correlazione" ed in definitiva applicano dei criteri analoghi al test dell'ipotesi nulla già menzionato.

FASE 2

Scegliere la espressione che meglio rappresenta la variabile "obiettivo" in funzione della variabile "causa" ritenuta significativa.

Questo corrisponde a tracciare una retta od una curva che interpolino, cioè che seguano l'andamento della nuvola dei dati.

Anche per questo esistono delle tecniche numeriche: noi dobbiamo prima stabilire che tipo di funzione vogliamo usare. Se questa è una retta, la sua espressione è $Y=AX+B$ in cui Y è la variabile "obbiettivo" ed X è la variabile "causa". Queste tecniche numeriche ci danno i valori di A e B che meglio interpolano i nostri dati e vanno sotto il nome di "Best Fit" o di analisi di "regressione".

Se invece abbiamo usato il procedimento grafico di tracciare ad occhio la retta o la curva interpolante, possiamo scriverne direttamente la espressione.

FASE 3

Se la nostra nuvola è molto stretta, la funzione che abbiamo trovato può essere considerata soddisfacente: gli scostamenti residui possono essere tollerati ed attribuiti a fattori che non vogliamo indagare e che consideriamo globalmente come "fattori casuali". Se invece gli scostamenti sono forti, possiamo calcolare per ciascun punto della nuvola gli scostamenti residui rispetto alla curva tracciata e tentare di spiegare questi scostamenti con una seconda variabile "causa". Ritorniamo quindi alla fase 1 e ripetiamo il procedimento: la nostra nuova variabile "obbiettivo" è lo scostamento residuo e la nuova variabile "causa" dovrà spiegare questi valori.

A quanto detto finora possiamo aggiungere alcune considerazioni:

— Quando, alla fase 1, scegliamo una variabile "causa", possiamo prendere una delle variabili elementari di cui disponiamo oppure una qualunque funzione di esse. Può essere ad esempio il quadrato o la radice quadrata di una delle variabili, oppure il prodotto di due delle variabili elementari che abbiamo a disposizione; in ogni caso avremo una serie di valori "causa" ed una serie di valori "obbiettivo" tra i quali tentare di stabilire la relazione.

— Abbiamo finora parlato genericamente di "variabili" causa, tuttavia, in senso stretto, il termine "variabile" si applica soltanto a quelle che presentano un valore numerico continuo. Ad esempio, la probabilità di una colonna secondo il pronostico, è, in questo senso, una variabile in quanto è espressa da un valore fra zero ed uno: maggiore è la probabilità, maggiore è il numero.

In altri casi, questo numero non esiste: ad esempio, se consideriamo fra le possibili cause la composizione della colonna, espressa dalla successione degli 1,X e 2, questa non è un numero e si parla più propriamente di "attributo" anziché di variabile. Naturalmente, quando abbiamo a che fare con "attributi", il solito diagramma può non avere senso, ma possiamo sempre rimediare con una tabella.

— Infine, è opportuna una considerazione su quante variabili potremo includere nel nostro modello. Questo dipende dal numero di dati che abbiamo a disposizione, ed abbiamo qui una trappola latente.

È chiaro che, per fare un caso limite, se abbiamo solo due dati possiamo scegliere qualunque variabile "causa" ed, adottando come funzione una retta, troveremo in ogni caso un modello perfetto. Questo modello perfetto andrà quasi sicuramente in pezzi appena proveremo ad applicarlo ad un nuovo terzo dato.

Si parla, a questo proposito, di "gradi di libertà":

— Il numero di dati di cui disponiamo nel nostro caso è 21, è un primo valore dei "gradi di libertà".

— Il numero di parametri che compaiono nel modello, per una retta sono 2, A e B, costituiscono i "gradi di libertà" del modello.

— La differenza sono i gradi di libertà dell'"errore".

Se i gradi di libertà dell'errore sono zero, il nostro modello risulterà perfetto ma assolutamente inattendibile. Occorre anche notare che, in un certo senso, anche tutte le variabili "di riserva", che abbiamo esaminato e scartato o che non abbiamo neppure preso in considerazione, dovrebbero entrare nei gradi di libertà del modello, quindi dobbiamo essere molto cauti. Se poi usiamo degli "attributi" la situazione è ancora più pericolosa, perché ogni singolo valore dell'attributo rappresenta un grado di libertà.

In definitiva, è una buona regola quella di accantonare una metà dei dati disponibili, scegliendoli a caso, di sviluppare il modello su una metà dei dati, lasciando comunque un congruo numero di gradi di libertà all'errore, di reinserire poi i dati accantonati a verificare la significatività del modello.

FASE 4

Il passo finale è infatti di verificare se il modello determinato è "significativo"; questa verifica procede secondo il solito schema a senso unico: supposto di aver trovato un certo allineamento fra i dati calcolati dal modello ed i dati effettivi, la verifica viene espressa da: "se i dati calcolati ed i dati effettivi sono indipendenti, cioè se non vi è fra loro nessuna relazione ed il nostro modello è quindi inefficiente, quale è la probabilità di ottenere "per caso" l'allineamento che abbiamo trovato?". Se questa probabilità è alta, non potremo respingere l'ipotesi di indipendenza dei dati e dovremo quindi considerare il nostro modello come fallito, in caso contrario, potremo concludere che l'ipotesi di indipendenza dei dati è "falsa" e considerare il modello come un primo risultato valido.

In pratica noi lavoriamo sempre per "falsificazione" di un modello. I dati non possono qualificare mai un modello come "vero' ma al massimo come "non manifestamente falso".

Se il modello supera la prova dei dati e se per altri motivi ci appare soddisfacente, possiamo tenerlo per buono ma in realtà non abbiamo motivi validi per considerarlo "vero": resta lì finché non subentri un modello migliore o finché nuovi dati sperimentali non lo dimostrino felicemente "falso" a sua volta.

Con un certo ottimismo si può considerare che un modello che ha su-

perato la prova dei dati si possa ritenere "corroborato"; certamente si sente "corroborato" il modellista.

La prima ipotesi considerata

La nostra prima analisi è fatta sui dati della figura 2 già esaminata prima. La prima ipotesi di strategia considerata è questa:

- Esistono due categorie di giocatori
  - Gli "spericolati", che puntano alla grossa vincita e giocano le colonne praticamente "a caso", distribuendosi fra tutte le colonne in modo praticamente uniforme.
  - I "sistemisti", che puntano alla vincita immediata e giocano a partire dalle colonne più probabili e spingendosi via via su quelle meno probabili, finché non hanno esaurito la somma stanziata.

Entrambi farebbero il nostro gioco, i primi "sovrapuntano" le colonne a bassa probabilità, i secondi "sovrapuntano" le colonne ad alta probabilità ed in mezzo dovrebbe rimanere una fascia "sottopuntata", corrispondente all'esaurimento delle risorse finanziarie dei sistemisti, dove ci possiamo sistemare noi.

Considerando però i dati della figura 2, la nostra ipotesi esce completamente distrutta, i dati stanno sopra e sotto alla retta di indifferenza in misura praticamente uguale sia per le colonne ad alta che a media che a bassa probabilità. Naturalmente potremmo fare una verifica quantitativa di questa constatazione, usando un metodo del tipo di quello già usato al capitolo precedente, ma l'esame visivo non incoraggia a perdere tempo su un'ipotesi così poco promettente.

Non ci sentivamo per niente corroborati.
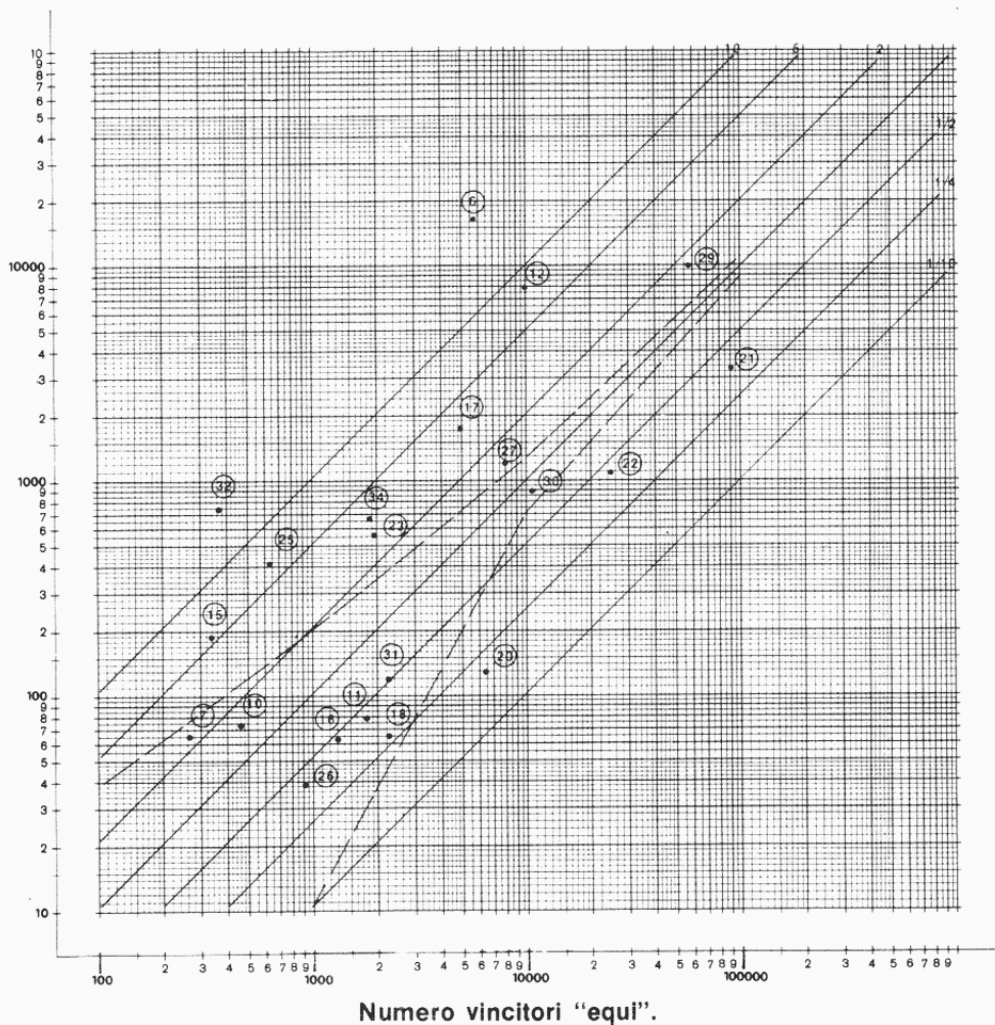
Figura 2 — Confronto fra vincitori "equi" ed effettivi.

Altre ipotesi alternative

La nostra ipotesi iniziale conteneva evidentemente degli errori che possiamo tentare di correggere.

Forse non esistono più giocatori con vecchie zie assolutamente ignoranti a cui far compilare una schedina realmente "a caso"; forse la categoria degli "spericolati" usa comunque almeno una trottola con una probabilità più ragionevole di 1,X e 2, oppure gioca sempre le stesse colonne fisse, il che ha praticamente lo stesso risultato: la frequenza delle giocate non dipende ugualmente dalla probabilità dei singoli risultati delle partite in modo diretto, però non si distribuisce uniformemente su tutte le colonne possibili, alcune di queste, quelle più palesemente improbabili, vengono scartate. Potremmo quindi chiamarli "trottolisti".

Anche per i sistemisti il ragionamento deve essere diverso: se facciamo l'ipotesi che la colonna più puntata sia sempre la prima in ordine di probabilità decrescenti, cioè quella a probabilità più alta, questo non significa che essa abbia in tutti i concorsi lo stesso valore di probabilità. In effetti esistono concorsi più o meno "piatti": in un concorso con molti pronostici "facili" la prima colonna ha un valore elevato di probabilità, che poi diminuisce rapidamente per le colonne successive, invece nei concorsi "piatti" anche la prima colonna ha un valore basso di probabilità e la diminuzione nelle colonne successive è meno veloce.

Ad esempio, in un concorso "ripido", come il N.31, la prima colonna in ordine di probabilità ha una probabilità di uscita di 540 per milione, che otteniamo moltiplicando le probabilità dei risultati più probabili di ciascuna partita, viceversa nel concorso n. 23, uno dei più "piatti", la probabilità della colonna più probabile è soltanto di 13 per milione.

È verosimile che in entrambi i concorsi la prima colonna abbia ottenuto lo stesso numero di puntate, mentre il rapporto delle probabilità è di oltre 40 a 1.

Modifichiamo quindi la nostra ipotesi per calcolare il numero di puntate dei "sistemisti"e consideriamo, non la probabilità della colonna, ma la sua "posizione" nella sequenza delle colonne dalla più alla meno probabile.

Abbiamo citato queste due varianti alla nostra prima ipotesi, perché sono quelle che entreranno nel modello finale, ma beninteso esistono o

possono esistere altri "meccanismi" di giocata più o meno importanti che intervengono nel determinare il risultato. Basta dare un'occhiata ai settimanali dedicati al totocalcio per trovare una quantità di suggerimenti ai sistemisti; tutti questi sono verosimilmente seguiti da una certa frazione degli scommettitori, per quanto farneticanti possano essere; il problema sta nel trovare "quanti" scommettitori seguono ciascuno di essi. Il modello che presenteremo è quindi largamente migliorabile, inserendovi un maggior numero di variabili.

C'è solo una ipotesi che, benché attraente non ci potremo forse mai permettere ed è quella di una variazione nel mix di strategie degli scommettitori dipendente dal numero di puntate totali.

In effetti, nei concorsi esaminati, si è passati da circa 120 milioni di puntate per il N.6 ad oltre 170 milioni e sarebbe ragionevole sospettare che questi 50 milioni di puntate aggiuntive possano non essere qualitativamente omogenei con i primi 120 milioni: potrebbero essere tutti nuovi "trottolisti", oppure tutti nuovi "sistemisti" oppure ancora essere gli stessi sistemisti di prima che "allungano" la puntata.

Tuttavia noi siamo in una situazione particolare: abbiamo per ogni concorso un campione molto numeroso che ci assicura una buona stabilità (campioni di 150 milioni di individui sono solo un sogno per chi non faccia il demografo in India o non si occupi di molecole) però il numero di concorsi che abbiamo è basso, nel nostro caso è di 21 ed anche prendendone un numero maggiore rimangono sempre pochi. Frazionare il totale delle puntate è una operazione troppo comoda, e per quanto possibile rischia di "spiegare tutto" introducendo gradi di libertà eccessivi nel nostro modello.

La strategia del sistemista

La nostra nuova ipotesi, per quanto riguarda i "sistemisti", è quindi che il numero di puntate raccolto da una colonna dipenda dalla posizione della colonna stessa nella successione ordinata delle colonne dalla più alla meno probabile.

Dovremo quindi determinare se la colonna vincente sia la prima, la centesima, la centomillesima o la milionesima della serie. Se finora a-

vremmo anche potuto fare a meno del calcolatore, ora questo ci è assolutamente indispensabile.

Anche con il calcolatore è comunque poco consigliabile avviarsi a produrre tutte le 1.594.323 colonne possibili, potrebbe richiedere intere giornate di calcolo; sommando tempi di calcolatore si riesce effettivamente a far su milioni a forza di millesimi; diversamente dalle lire, checché ne dicano i promotori della giornata del risparmio.

Useremo quindi un metodo chiamato metodo "Montecarlo". È meglio precisare subito, per non aizzare di più i fautori del risparmio, che non si tratta di un metodo per arricchirci rapidamente, ma soltanto di una tecnica numerica.

Prenderemo in esame solo un campione della popolazione totale di 1.594.323 colonne. Questo campione deve essere costituito da colonne distribuite uniformemente entro la popolazione ed a questo scopo noi genereremo queste colonne del campione "a caso", assegnando uguali probabilità ai segni 1, X e 2.

Questo procedimento, proprio in quanto utilizza una casualità nella composizione del campione, prende il nome da Montecarlo, la città santa del "caso".

Per ciascuna delle colonne generate, calcoleremo la probabilità sulla base del pronostico relativo al concorso che ci interessa e metteremo queste colonne in ordine di probabilità decrescente.

Supponiamo di aver generato 1594 colonne, cioè circa un millesimo del totale e di aver calcolato per ciascuna di queste la probabilità, che indichiamo con P1, P2, P3,,,, P1594.

Se prendiamo la prima colonna della serie, quella a probabilità più alta, P1, possiamo dire che questa rappresenta le prime mille colonne della popolazione, che queste avranno una probabilità media uguale alla probabilità P1 ed una probabilità inferiore alla media fra P1 e P2.

Proseguendo a questo modo per tutte le 1594 colonne generate arriveremo a costruire, ovviamente nel calcolatore, una tabella del tipo della figura 3. A questo punto diventa immediato il determinare in quale posizione si collochi una qualsiasi colonna, basta calcolarne la probabilità e vedere nella tabella in quale intervallo venga a posizionarsi interpolando fra i valori limite.

COLONNE DA: A:	PROBABILITÀ LIMITE
1	P(MAX) (1)
2 – 1000	(P1+P2)/2
1000 – 2000	(P2+P3)/2
2000 – 3000	(P3+P4)/2
1593000 – 1594000	(P1594+P(MIN))/2
1594323	P(MIN) (2)

(1) Probabilità della colonna più probabile.
(2) Probabilità della colonna meno probabile.

Figura 3

Naturalmente questo procedimento è intrinsecamente approssimato, benchè esistano raffinamenti allo schema semplificato descritto, ma se generiamo un numero di colonne abbastanza grande il risultato risulta sufficientemente preciso: 1594 colonne sono già un discreto numero.

In allegato è riportato il programma relativo che può spiegare, con un linguaggio più preciso, il funzionamento del metodo.

Applicando questo procedimento ai nostri 21 concorsi, otteniamo in definitiva la tabella in figura 4 che riporta la posizione della colonna vincente di ciascun concorso nella serie relativa allo stesso concorso dalla colonna 1, la più probabile, alla 1.594.323, la meno probabile.

Nella stessa tabella è riportata, per futuro uso, la struttura della colonna vincente, espressa dal numero di X e di 2 presenti ed il numero dei vincitori effettivi del concorso. Il numero dei vincitori effettivi è "normalizzato" ad un concorso standard da 150.000.000 colonne giocate.

Concorso N.	Posizione Colonna Vincente	Struttura	Vincitori (150.000.000 giocate)
6	52.000	6/X	20.100
7	135.000	4/X—3/2	78
10	530.000	5/X—3/2	82
11	230.000	4/X—2/2	84
12	94.000	5/X	8.500
15	1.100.000	5/X—1/2	185
16	280.000	6/X—3/2	63
17	76.000	4/X—2/2	1.750
18	195.000	7/X—2/2	65
20	44.000	6/X—3/2	140
21	620	6/X—2/2	2.950
22	10.500	6/X—1/2	1.050
23	295.000	5/X—3/2	520
25	680.000	4/X—1/2	405
26	410.000	1/X—5/2	36
27	41.000	8/X	1.020
29	330	7/X—1/2	9.500
30	30.000	6/X—2/2	850
31	145.000	5/X—2/2	115
32	470.000	4/X-3/2	720
34	190.000	3/X-2/2	710

Figura 4

Possiamo a questo punto, riportarci i dati in diagramma e verificare come si presentano: il diagramma è riportato in figura 5, sempre utilizzando una scala logaritmica, nonostante la quale il concorso N. 15 esce leggermente dal quadro.

Questo diagramma è decisamente più "corroborante" di quello della figura 2. L'andamento decrescente del numero di puntate con la posizione progressiva della colonna sembra abbastanza evidente: intorno
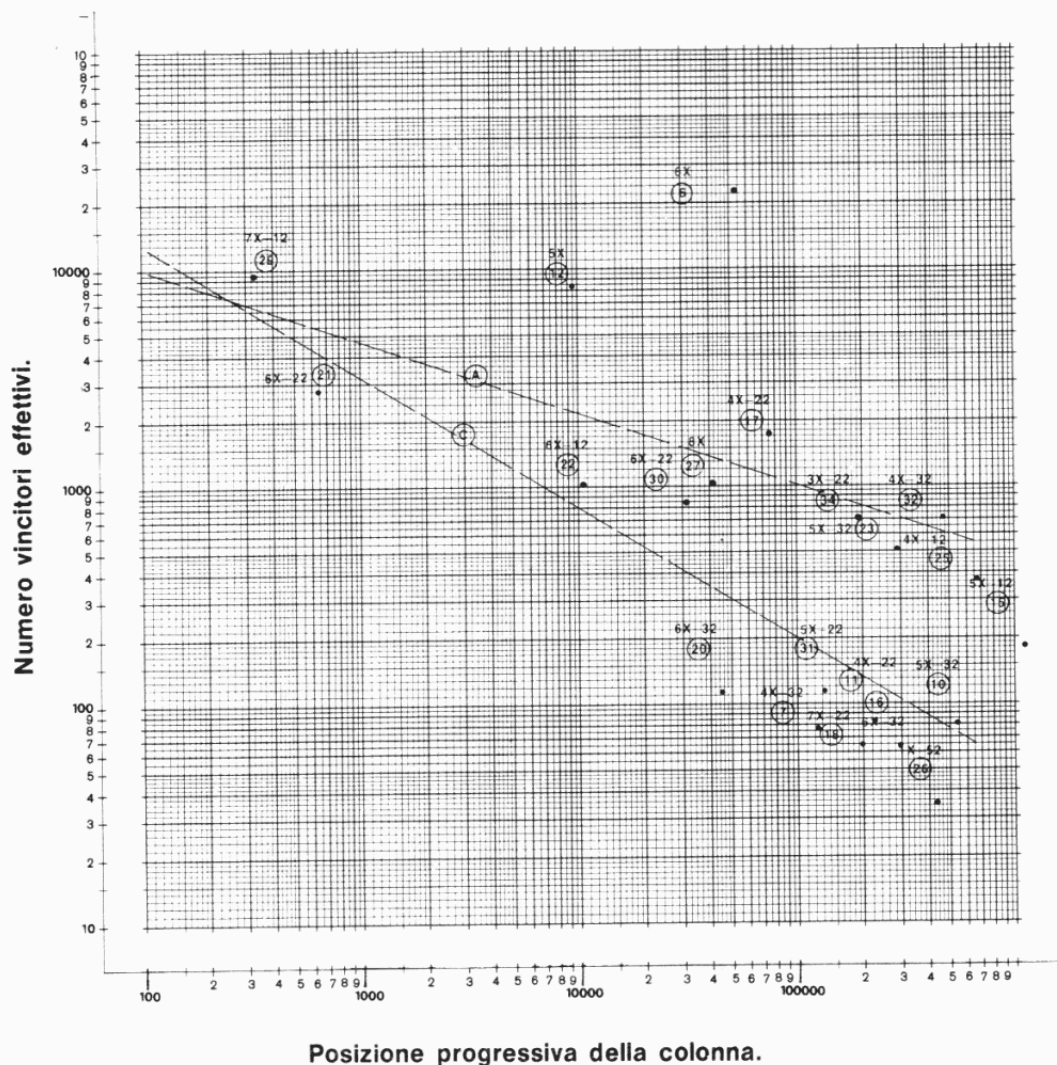
Figura 5 — Normalizzata a 150.000.000 giocate.

ad una retta che tirassimo fra i concorsi 29-21 ed i 25-32 c'è un buon allineamento di punti (retta A) rimarrebbero fuori i concorsi 6-12 da un lato e la nuvola dei concorsi in basso, però esaminando la struttura delle colonne vincenti, notiamo che i concorsi 6-12 non hanno nessun 2, mentre la nuvola dei concorsi in basso ne presenta parecchi: abbiamo buone speranze che queste anomalie residue vengano spiegate con lo schema della "trottola".

La strategia del "trottolista"

La strategia che attribuiamo al trottolista è più semplice: compila la schedina senza guardare quali squadre giochino in ciascuna partita, ma facendo semplicemente girare una trottola poligonale che porta sulle facce i simboli 1, X e 2 in proporzione diversa ed inserisce nella colonna il risultato che esce sulla trottola. Può essere una trottola a sei facce con tre segni 1, due segni X ed un segno 2, oppure a dieci facce con cinque segni 1, tre segni X e due segni 2, oppure dei dadi o qualcosa del genere.

Questo è un metodo materiale effettivamente usato da alcuni giocatori, non sappiamo se tanti o pochi, ma è anche un modello ideale che può rappresentare parecchi altri comportamenti che danno luogo comunque allo stesso risultato.

Producono lo stesso risultato, ad esempio:

- Chi gioca sempre certe colonne fisse, scelte fra le colonne che hanno una combinazione di 1–X–2 ritenuta più probabile.
- Chi elimina da un certo sistema le colonne che viceversa hanno una combinazione di 1–X–2 ritenuta poco probabile.

In definitiva tutti quelli che introducono nel loro modo di giocare delle considerazioni sulla struttura della colonna senza considerare le partite effettive.

Queste sono considerazioni palesemente assurde, perché sopprimono quella che è l'unica informazione valida al Totocalcio, e cioè che, poco o tanto, una squadra è più forte di un'altra e quindi avrà più facilmente un risultato favorevole, tuttavia se il nostro avversario le usa, ci interessa sapere come lo fa.

In un modo più o meno esplicito, dietro a questi comportamenti c'è la convinzione comune "all'uomo-della-strada ed al filosofo K. Marbe" che dopo l'uscita di un risultato diventi più probabile l'uscita di un risultato diverso e che quindi il numero di segni uguali sulla schedina non possa superare un certo numero e si ritiene quindi di migliorare un pronostico "tecnico", basato sulla probabilità dei singoli risultati, dimenticando le
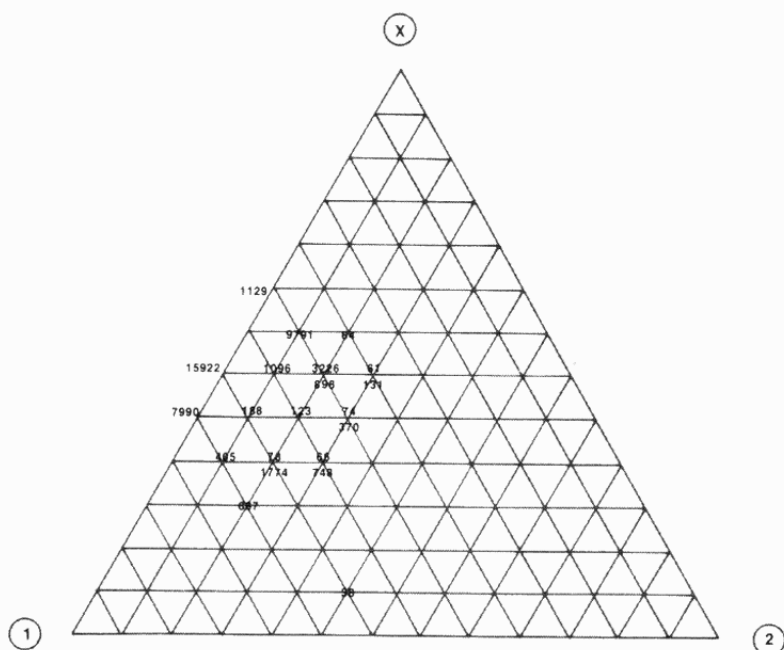
Figura 6 — Configurazioni vincenti e numero giocate.

singole partite e ritoccando la struttura globale della colonna.

In realtà il risultato è esattamente opposto, il pronostico viene peggiorato e non migliorato, perché non esiste una "memoria dentro la schedina", i segni 2 non si consumano man mano che escono.

La strategia "trottola" può essere efficacemente analizzata con l'aiuto di diagrammi del tipo della figura 6, che danno una buona veduta di insieme. Il diagramma è un triangolo equilatero con i lati divisi in 13 parti quante sono le partite di una colonna, ogni punto di incrocio rappresenta una possibile configurazione di colonna, composta da un determinato numero di 1,2,X.

L'ordine non è importante, conta solo il numero totale di ciascun segno. I vertici del triangolo contrassegnati con 1,X,2 rappresentano le colonne costituite rispettivamente da 13 segni uguali 1,X,2.

Un punto di incrocio interno qualsiasi rappresenta una colonna con tanti segni 1 quanto è la distanza dalla base opposta al vertice 1, e analogamente per i segni X e 2.

È più semplice provare sul disegno, che tentare di seguire un discorso di spiegazione più lungo, o di farlo.

Nella figura 6 sono riportati i 21 concorsi presi in esame con il relativo numero di vincitori. Alcune configurazioni sono uscite due volte e sono riportati i due valori del numero di vincitori.

Questo diagramma spiega un altro aspetto della strategia del trottolista, che stranamente in certi casi predilige e in altri casi aborre i risultati già usciti.

La strategia del trottolista, in effetti, è pura fantasia, al confronto il "sistemista" è un mostro di razionalità: si ferma quando non ha più soldi, e non si vede chi potrebbe fare diversamente.

Comunque, il trottolista, di fronte al diagramma di figura 6 dice: "ecco, vedi bene anche tu che certi risultati escono, anche più volte durante l'anno ed altri non escono mai. Quindi io gioco quelli che escono".

In realtà questi risultati si giustificano perfettamente sul piano tecnico ma non rappresentano affatto un buon motivo per prediligere certe strutture di colonna in luogo di altre.

La giustificazione tecnica si trova considerando i diagrammi delle figure 7,8,9,10.

Nelle figure 7 ed 8 sono riportate, per due diversi tipi di "trottole", rispettivamente una 50%:1/30%:X/20%:2 ed una 50%:1/45%:X/5%:2, le probabilità di generare una colonna della configurazione specifica indicata.

Il campionato idealmente, è anch'esso una "trottola" che produce risultati secondo una probabilità che è vicina alla 50/30/20, in effetti nei 21 concorsi presi in esame le percentuali sono state 49/36/15.

Quindi, se proprio vogliamo ignorare le partite specifiche che vengono giocate perché ci sentiamo incapaci di fare un pronostico tecnicamente valido, dovremmo riconoscere alle diverse colonne la probabilità che risulta dalla figura 7, che attribuisce ad esempio alla colonna 13/1 la probabilità di 122 per milione ed alla colonna 7/1–6/X una probabilità molto inferiore di soltanto 5.7 per milione.

Ma, dice il "trottolista", la colonna 13/1 non è mai uscita, mentre altre escono continuamente.
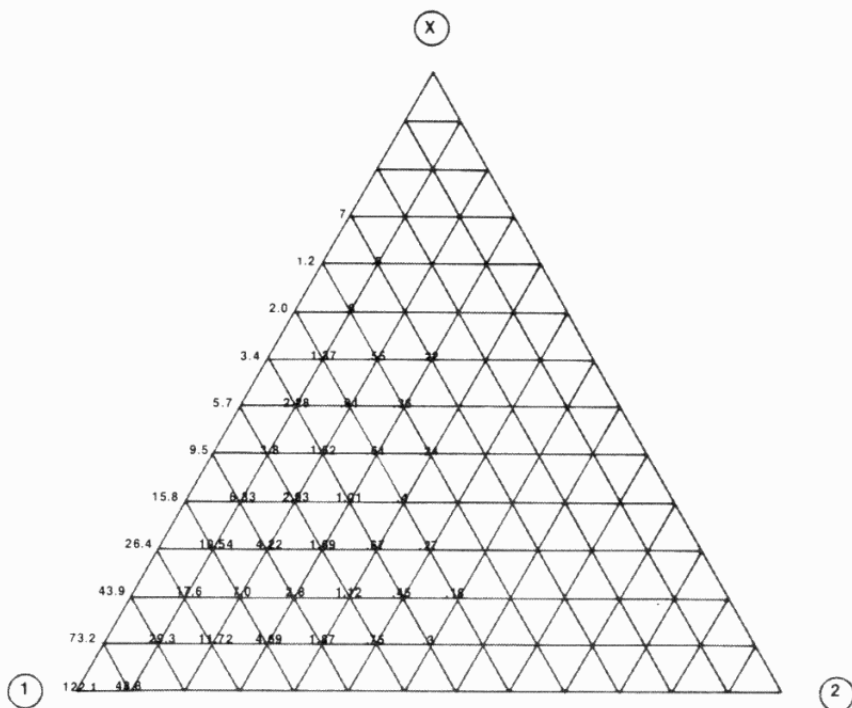
Figura 7 — Probabilità secondo schema "Trottola" 50-30-20 (i numeri sono/per milione).
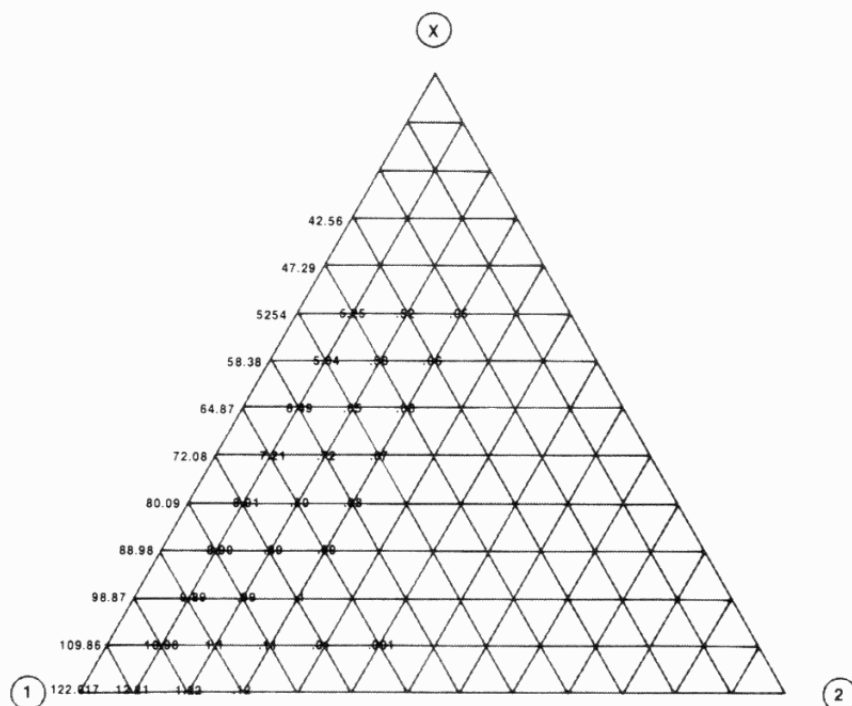
Figura 8 — Frequenze sistema "Trottola" 50-45-5 numeri per milioni.
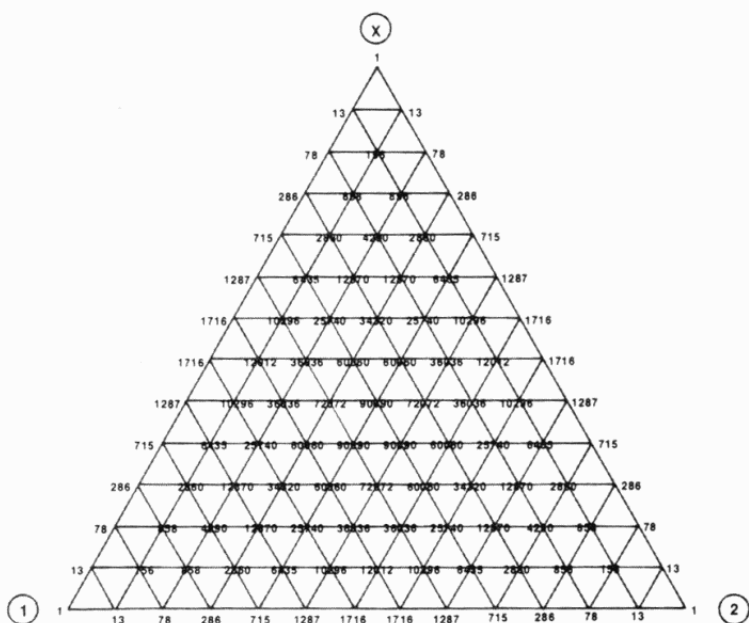
Figura 9 — Schede totali combinatorie.
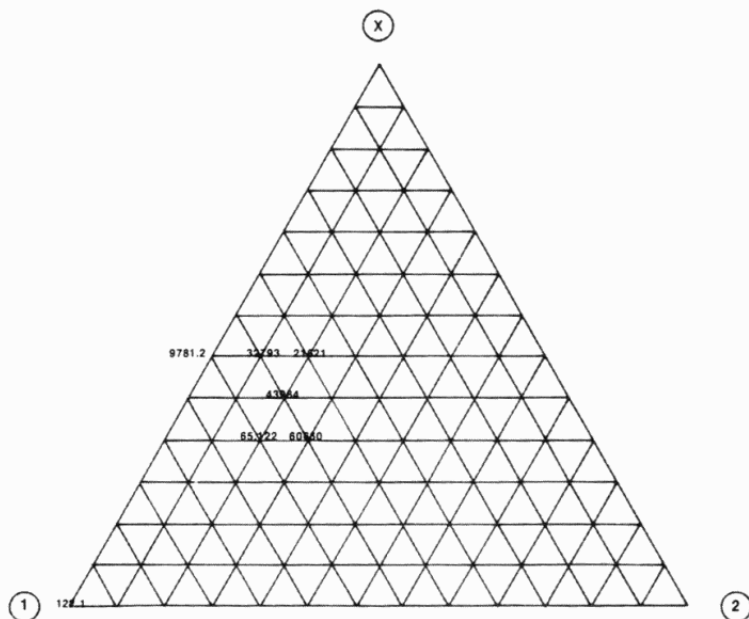
Figura 10 - Probabilità di uscita trottola 50-30-20 (le probabilità sono per milione).

La spiegazione è molto semplice, esiste una sola colonna con 13 segni 1, mentre esistono, ad esempio 1716 modi diversi di formare una colonna con 6 segni X e 7 segni 1. Nella figura 9, è riportato appunto il numero di colonne di ciascun tipo possibili.

La figura 10 sintetizza i dati delle figure 7 e 9: moltiplicando le probabilità della figura 7, che rappresenta la probabilità di produrre "una" colonna di un certo tipo, per il numero di colonne possibili di quel tipo, preso dalla figura 9, otteniamo la probabilità globale che quel tipo di colonna si verifichi. Alcuni di questi risultati sono riportati nella figura 10, da cui vediamo ad esempio che la probabilità che si verifichi "la" colonna 13/1 è sempre di 122 per milione, mentre la probabilità che si verifichi "una qualsiasi" delle colonne 4/X–2/2 è di 65.122 per milione. Tutto questo però è di scarso aiuto al giocatore, che di tutte le 25740 colonne del tipo 4/X–2/2 ne gioca una sola.

La sintesi delle due strategie "sistemisti" e "trottolisti"

A questo punto dobbiamo tentare una sintesi delle strategie seguite dai due gruppi che abbiamo ipotizzato e vedere se riusciamo a calcolare il numero totale di giocatori per colonna.

Avevamo ipotizzato che il numero dei vincitori di un concorso fosse costituito da una certa frazione di giocatori "sistemisti" che danno luogo ad un numero di puntate che dipende dalla posizione della colonna e da un certo numero di "trottolisti" le cui puntate dipendono dalla probabilità di generare la colonna vincente con una trottola opportuna.

Purtroppo anche questa ipotesi di "somma" delle due frazioni non regge alla prova dei dati:

Rispetto alla retta (A) tracciata nella figura 5, che abbiamo supposto potesse rappresentare la strategia dei "sistemisti", abbiamo circa 6000 puntate in più per il concorso 12 e circa 20.000 in più per il concorso 6 da attribuire ai "trottolisti".

Questo non ci torna con le probabilità della figura 7 che ci giustificano al massimo un migliaio di puntate. Potremmo supporre che la trottola sia del tipo di figura 8, che alzando la probabilità del segno X fino al massimo che possiamo praticamente immaginare, ci alza la probabilità per le colonne che non contengono segni 2, e quindi aumenta i "trottoli-

sti" per i concorsi che, come il 6 ed il 12, presentano solo 1 ed X; anche questo però è insufficente. Inoltre, la retta (A) che abbiamo tracciato tanto per dare un'idea dell'allineamento, è sicuramente troppo alta: una retta che ci dia la frequenza di giocata dei "sistemisti" deve stare tutta al di sotto dei punti, se poi vogliamo aggiungerci almeno qualche "trottolista"; questo aumenta ancora il numero di "trottolisti" per i concorsi 6 e 12 ad un livello tale che non possiamo sperare di giustificarli con questo schema.

Stavamo per abbandonare l'impresa che, cominciata per gioco sembrava diventare troppo lunga quando Totomac, con uno dei suoi soliti colpi anfibi, produsse il modello riportato al capitolo seguente, che ha superato la prova dei dati.

Gli abbiamo chiesto come fosse arrivato a definire quel modello, ma ci rispose che non lo sapeva, che era un'idea che gli era venuta per caso, il che non stupisce conoscendo i suoi processi mentali.

Gli abbiamo chiesto quale meccanismo seguissero i giocatori per produrre il risultato e rispose che non sapeva nemmeno quello ma acconsentì a spiegarsi un po' più diffusamente: non ci sono diverse categorie di giocatori, ce n'è una sola, o almeno le altre sono trascurabili perché rappresentano poche puntate. L'unica categoria importante è quella dei "sistemisti" che giocano come avevamo supposto con frequenze decrescenti secondo la posizione della colonna, però correggono la puntata in base ad uno schema "trottola", quindi non dobbiamo sommare le due strategie ma moltiplicarle fra loro".

Vedendoci perplessi, aggiunse con sano empirismo pre-newtoniano: "dopo tutto gli epicicli di Tolomeo sono andati bene per anni".

CAPITOLO 3

IL MODELLO DEFINITIVO

A questo punto abbiamo definito un modello della strategia globale dei giocatori. Questo ci permette di calcolare il numero di giocate che riceverà una qualunque fra le possibili colonne in base a:

— pronostico in forma standard probabilistica per il concorso in esame

— struttura della colonna

Noi ipotizziamo l'esistenza di una sola categoria rilevante di giocatori: i sistemisti. Ciascuno di loro gioca secondo una strategia base che ha due componenti:

— giocare le colonne a partire dalla più probabile e spingendosi in avanti verso le meno probabili, fino ad esaurimento della somma stanziata;

— scartare le colonne che presentano una struttura "sgradita". Il "gradimento" di una colonna è espresso dalla probabilità di una trottola.

Siccome non tutti i sistemisti stanziano la stessa somma, non si fermeranno tutti allo stesso punto, in pratica la frequenza di giocata andrà quindi gradualmente diminuendo man mano che ci si allontana dalla prima colonna, con un andamento che dipende da un fattore economico, le tasche degli scommettitori, e non la probabilità di uscita della colonna.

La curva adottata per rappresentare la prima componente della strategia è rappresentata in figura 5 dalla retta indicata con (C). La figura 5 include solo l'intervallo fra la centesima e la milionesima colonna della sequenza, tuttavia ovviamente la retta si estende in tutto l'intervallo dalla colonna 1 alla colonna 1.594.323. La rappresentazione analitica di

questa curva per un concorso di 150.000.000 di giocate è, per chi ci sia interessato:

$$\frac{206.000}{NG^{06}}$$

In cui NG rappresenta la posizione progressiva della colonna.

Chi sia interessato solo alla applicazione del metodo, non ha comunque motivo di preoccuparsi della formula: i valori della funzione sono già stati calcolati ed inseriti nel modello in forma tabellare.

Questa funzione dice che: la prima colonna della serie, la più probabile riceve 206.000 puntate su 150.000.000 di colonne giocate, la 100-esima, 1000-esima, 10.000-esima, 100.000-esima, 1.000.000-esima ne ricevono rispettivamente 13000, 3200, 780, 195, 48; questi dati li leggiamo approssimativamente dal grafico in figura 5. L'ultima colonna, in posizione 1.594.323 ne riceve circa 39.

Dobbiamo ora definire la seconda componente della strategia, quella che corregge i risultati della prima introducendo una intensificazione od una rarefazione delle puntate sulle colonne secondo la loro struttura.

Risulta abbastanza significativa a questo fine la probabilità di generazione di una struttura data con una trottola 50/30/20; alcuni di questi dati di probabilità sono riportati nella figura 7, gli altri si possono calcolare facilmente.

Questa probabilità, espressa per milione, viene moltiplicata per il numero di giocate dato dalla prima componente, se la probabilità è maggiore di 1 per milione, la frequenza di giocata sulla colonna viene intensificata, se è minore di 1 viene ridotta.

La formula definitivamente adottata, indicando con P7 per brevità la probabilità di figura 7, è quindi:

206.000 * P7 / NG↑0.6

Questa formula contiene, esplicitamente o no, diverse costanti:

- Il 206.000 come punto di partenza.
- Lo 0.6 come esponente di NG.
- La trottola 50/30/20, un'altra trottola potrebbe anche andare meglio.

— L'esponente di P7 che è implicitamente uguale ad 1, ma che non si vede perché debba essere proprio così.

Queste costanti sono state adottate in via di prima approssimazione, ma una verifica di congruenza ed un affinamento possono portare ad un miglioramento del modello.

La verifica di congruenza dovrebbe verificare che effettivamente questa formula produca 150.000.000 di giocate.

L'affinamento dei coefficienti costanti dovrebbe verificare quali siano i valori che forniscono il miglior accordo fra valori calcolati e valori effettivi.

Entrambe queste verifiche sono state fatte in modo piuttosto approssimativo, il motivo è che il "gruppo di lavoro" era già passabilmente soddisfatto del risultato raggiunto.

Il risultato finale è sintetizzato nella tabella in figura 11. In questa tabella sono riportati, per ogni concorso e sempre per 150.000.000 di giocate:

— La posizione progressiva e la struttura della colonna vincente.

— Il numero di vincitori secondo la curva (C) e la probabilità secondo la trottola 50/30/20.

— Il numero di vincitori calcolato, l'effettivo e l'"equo" che corrisponde alla probabilità della colonna secondo il pronostico (quest'ultimo dato è uguale a quello della figura 2 salvo la "normalizzazione" a 150.000.000 giocate).

Dall'esame visivo della tabella, risulta che la nostra formula è sufficientemente precisa da discriminare fra le colonne "sovrapuntate" e le colonne "sottopuntate" che è l'obbiettivo che ci interessa; in effetti, nella maggior parte dei casi, quando il calcolo ci fa prevedere una colonna "cattiva", con un numero di puntate superiore al valore "equo", questa risulta "sovrapuntata" anche come numero effettivo di vincitori e viceversa, quando ci fa prevedere una colonna "buona", con un numero di vincitori calcolati inferiore al valore "equo", questa risulta "sottopuntata" anche come numero effettivo di vincitori.

L'osservazione visiva non è però sufficente, l'entusiasmo del ricercatore può portare a sopravvalutare un risultato che, dopo tutto, potrebbe anche essere stato un caso. Abbiamo però a disposizione, anche qui, una metodologia quantitativa analoga a quella usata al capitolo 1.

Supponiamo di classificare i nostri 21 concorsi nel modo seguente:

- "Calcolato buono" se i vincitori calcolati sono meno dell'80% dei vincitori equi.
- "Calcolato cattivo" se i vincitori calcolati sono oltre il 120% dei vincitori equi.
- "Calcolato indifferente" se i vincitori calcolati stanno nei limiti 80-120% dei vincitori equi.

Riclassifichiamo ancora i concorsi come:

- "Effettivamente buono" se i vincitori effettivi sono meno dell'80% dei vincitori equi.
- "Effettivamente cattivo" se i vincitori effettivi sono oltre il 120% dei vincitori equi.
- "Effettivamente indifferente" se i vincitori effettivi stanno nei limiti 80-120% dei vincitori equi.

La figura 12 è una "tavola di contingenza", che riporta il numero di concorsi che cadono in ciascuna casella corrispondente all'incrocio delle classi "calcolato" ed "effettivo". (C) Rappresenta il valore cacolato, (E) rappresenta il valore effettivo e (P) rappresenta il valore equo. Il test di significatività, il programma è in allegato, ci permette di concludere che esiste una relazione significativa fra i valori calcolati ed i valori effettivi.

E quindi che il nostro modello "funziona".

N.B. I concorsi contrassegnati con * sono quelli che risultano, nelle rispettive colonne, come "calcolati buoni" ed "effettivamente buoni" rispetto al valore "equo" 20%.

N.	Strutt.	Posiz. Progress.	Vincit. Curva (C)	Trott. 5/3/2	Vincitori Calcolati	Vincitori Effettivi	Vincit. Equi
6	6X	52.000	300	5.7	1.710	20.120	713
7	4X/32	135.000	170	1.01	172	78	30
10	5X—32	530.000	70	.61	43	82	50
11	4X—22	230.000	120	2.53	303	* 84	166
12	5X	9.400	800	9.5	7.600	8.500	1.020
15	5X—12	1.100.000	48	3.8	182	185	32
16	6X—32	280.000	105	.36	* 38	* 63	114
17	4X—22	76.000	240	2.53	607	1.750	472
18	7X—22	195.000	130	.55	* 71	* 65	193
20	6X—32	44.000	320	.36	* 115	* 140	620
21	6X—22	620	4.000	.91	* 3.640	* 2.950	9.500
22	6X—12	10.500	750	2.28	* 1.710	* 1.050	2.550
23	5X—32	295.000	102	.61	* 62	520	162
25	4X—12	680.000	60	6.33	380	405	63
26	1X—52	410.000	85	.75	* 45	* 36	83
27	8X	41.000	340	2.0	680	1.020	685
29	7X—12	330	5.500	1.37	7.535	9.500	5.500
30	6X—22	30.000	400	.91	* 364	850	972
31	5X—22	145.000	150	1.52	228	* 115	192
32	4X—32	470.000	80	1.01	81	720	34
34	3X—22	190.000	135	4.22	570	710	186

Figura 11 Tabella riassuntiva

	EFFETTIVI		
	E < 9.8*P		E > 1.2*P
C < 0.8*P	6	1	1
CALCOLATI	1	0	2
C > 1.2*P	1	0	9

Figura 12 Tavola di contingenza

Osservazione sul modello definitivo

Una obbiezione comune al sistema Totomac, da parte di chi non ne abbia capito il funzionamento, è questa: "questo sistema mi farebbe giocare delle colonne poco giocate dagli altri, quindi improbabili, quindi anche se la vincita attesa è maggiore, la probabilità di vincita è bassa e in definitiva come "sistema" non è quanto di più desiderabile si possa immaginare".

Questa obbiezione non è giustificata, perché il Totomac si propone di individuare le colonne poco giocate, sia fra quelle ad alta probabilità che fra quelle a bassa probabilità, e quindi non produce "necessariamente" colonne a bassa probabilità.

Tuttavia, anche chi ha capito perfettamente il sistema, potrebbe sollevare un legittimo dubbio: "d'accordo, teoricamente il Totomac mi può dare sia colonne ad alta che a bassa probabilità, purché risultino poco giocate, però temo che, in pratica, finirà col darmi solo colonne poco probabili perché quelle ad alta probabilità risulteranno "sovrapuntate" e quindi me le farebbe scartare".

Questa obbiezione non può essere respinta tout-court, ma abbiamo il conforto di una verifica sperimentale che questo non succede: gli scommettitori, per nostra fortuna, "sottopuntano" anche colonne a probabilità molto alta. Dalla tabella in figura 11, risulta che prendendo in esame i concorsi in cui sono uscite le colonne a più alta probabilità, quelli con un numero di vincitori "equi" oltre il migliaio, due su quattro sono risultati "sottopuntati", tra cui il concorso 21 che è quello con numero di vincitori equi più alto di tutti.

Viceversa, prendendo in esame i cinque concorsi in cui sono uscite le colonne a più bassa probabilità, quelli con vincitori "equi" sotto a 80, tutti questi sono risultati "sovrapuntati".

Ci sembra di poter concludere che, almeno per quanto possiamo vedere da questi dati, questo timore non sia confermato, anzi che succeda proprio il contrario.

CAPITOLO 4

COMMENTI VARI

Girando oziosamente fra le dita il bicchiere di scotch, Baccarat chiese: "ma il Totocalcio cosa dirà di questo sistema?".

"Assolutamente niente" – rispose Van Random – "È un sistema come un altro e ciascuno è padrone e libero di giocare come vuole e come sa".

"Ma se tutti si mettono a giocare in questo modo?" – Baccarat proseguiva nelle sue riflessioni.

A Van Random bastavano non più di cinque secondi per impostare la trattazione completa del problema: "allora il Totocalcio diventa un gioco 'equo' – rispose – questo sistema tende a portare la frequenza di giocata su tutte le colonne ad un valore che sta in una proporzione costante con la loro probabilità di uscita. Quindi realizza proprio le condizioni di 'equità'. Istituzionalmente quindi, per tornare alla tua prima domanda, il Totocalcio dovrebbe essere contentissimo che queste condizioni si realizzino".

"Possiamo anzi dire" – proseguì spaziando olimpicamente – "che questo sistema è IL SISTEMA perché è un 'Minimax' nel senso della teoria dei giochi. Ogni giocatore è un 'aggressore' nei confronti dell'insieme degli altri giocatori e contemporaneamente contribuisce a produrre una strategia 'difensiva' collettiva che contrasta l'attacco del singolo avversario. Ora come ora, l'insieme delle strategie 'offensive' degli scommettitori non produce una strategia difensiva globale efficace, l'abbiamo visto sperimentalmente. Questo sistema sfrutta in effetti questa situazione, però, tendendo a costituire una condizione di 'equità', produce anche automaticamente la strategia difensiva ottima, quindi è stabile".

"D'altronde" – stava quasi sdraiato in poltrona e divagava a ruota libera – "tutto questo al totocalcio lo sanno certamente benissimo. È il loro mestiere e avranno in casa almeno venti persone che il problema lo conoscono perfettamente. Anche perché in questo sistema non c'è poi

un gran che di originale, come teoria. L'unica cosa nuova, e forse neanche quella, è di aver trovato che con il solo dato della colonna vincente e del numero di giocate su di essa è possibile identificare la strategia globale degli scommettitori. Noi finora abbiamo un modello molto approssimativo, ma sicuramente si può fare di meglio, lavorandoci su ancora un po'".

"Pensa poi cosa potremmo fare" — Van Random fantasticava pigramente, volute di numeri sembravano ondeggiare mollemente intorno al suo cranio — "se potessimo avere anche le colonne che vincono con il dodici e il numero di puntate su ciascuna di esse, dopotutto sono solo altre 26 colonne per settimana, ma la nostra informazione si moltiplicherebbe di altrettante volte e potremmo lavorare su basi molto più sicure".

Baccarat scattò come un grillo: "e perché non ce le facciamo dare?". "Ma non te le danno mica" — rispose Van Random — "questo è oro puro, amico mio. Prima ho detto che sarebbe bello avere i dodici, ma pensa se le avessimo tutte. Tutte le 1.594.323 colonne in bell'ordine, con il loro numero di giocate a fianco. Nel giro di tre o quattro settimane credo che potremmo sapere assolutamente tutto. Anche se c'è una vecchietta che gioca sempre una colonna di due fissi sperando nei miliardi. Allora fare un modello stravincente sarebbe un gioco da ragazzi".

"Wow" — abbaiò il bassotto entusiasmato — "gli conteremmo i peli uno per uno, agli altri scommettitori". Da quando è stato ribattezzato Bernoulli, il bassotto ogni tanto partecipa alla conversazione, e quasi sempre con un tono di trionfalismo pecoreccio da neo-deputato di maggioranza che dà molto sui nervi a tutti.

"Ma chi ti dice che non si possono avere?". Insisteva Baccarat.

"Ma perché non se ne parla neanche" — tagliò corto Van Random — "queste cose o sono di dominio pubblico, oppure non deve saperle nessuno. I dati devono essere protetti in modo assoluto, prima ovviamente, ma anche dopo il concorso. Altrimenti sarebbe come truccare una roulette al Casinò. Sarebbe da galera".

"Quanti anni dici che mi darebbero?" — Baccarat non si dava per vinto. "Cosa vuoi che ne sappia di quanti anni ti darebbero" — Van Random era veramente spazientito ed allarmato dall'atmosfera progettuale della assemblea — "tra l'altro" — cercò di minimizzare — "probabilmen-

te non lo sanno nemmeno al totocalcio, te l'immagini trascrivere 150 milioni di schedine. Sarà tanto se l'hanno fatto un paio di volte, come studio, in tutta la loro esistenza".

"Già, ma adesso è meccanizzato, ci sarebbe tutto pronto" – anche Totomac si schierava con il partito di attacco.

"Ma insomma, è impensabile" – Van Random vedeva insidiata la purezza del 'caso' e si batteva con tutti i mezzi a denti stretti – "anche con il calcolatore, sono 150 milioni e passa di colonne con tredici caratteri ciascuna, sono anche fisicamente intrasportabili".

"Mica tanto" – dovetti precisare – "con meno di 5 megabytes ti porterei a casa tutto. Le colonne non ti servono, una colonna, vista come un numero in base tre è solo un numero progressivo da 0 a 1.594.322. Una diversa colonna, un diverso numero. Ti basta il numero di giocate per colonna messo in ordine; tre bytes ciascuno sono più che sufficienti.

Poi naturalmente ti ci vuole una macchina robusta, ma varrebbe la pena.

"Allora provate a chiederglieli e vedrete cosa vi rispondono" – ribattè Van Random – "comunque ho capito e non vi conosco più. Siete tutti una banda di 'Computer criminals' in potenza, compreso quello lì" – concluse indicando il bassotto.

Si alzò e se ne andò.

PARTE SECONDA

IL PRONOSTICO PROBABILISTICO

CAPITOLO 5

USO DEL PRONOSTICO NEL SISTEMA TOTOMAC

Nel nostro sistema, come abbiamo visto, viene usato un pronostico in forma probabilistica, che assegna a ciascuno dei tre possibili risultati di ciascuna partita una probabilità percentuale. La somma delle probabilità dei tre risultati è ovviamente cento. È naturalmente la stessa cosa dividere tutto per cento ed avere probabilità espresse da numeri con decimali e somma totale uguale ad uno.

Nella parte prima, abbiamo usato un pronostico, quello della "Repubblica" che essendo pubblicato prima delle partite, ci assicurava di non essere inquinato, anche involontariamente, da risultati noti.

Infatti, lavorando a posteriori, se avessimo formulato un pronostico nostro sarebbe stato difficile, nonostante la miglior buona volontà, non farsi influenzare da quello che era in realtà accaduto.

Nello sviluppo del sistema abbiamo usato questo solo pronostico, ma in realtà il sistema stesso lo usa due volte, con scopi diversi, e, nell'uso pratico potrebbe essere conveniente differenziare le due funzioni ed usare due pronostici diversi, ciascuno per lo scopo che ci interessa.

Il primo uso del pronostico

Il primo uso del pronostico è per determinare il numero di puntate che riceverà una determinata colonna da parte degli altri giocatori.

Questo viene fatto quando noi usiamo il pronostico per determinare la posizione della colonna nella successione a probabilità decrescente di tutte le possibili colonne.

A questo fine, il miglior pronostico che possiamo usare è quello che corrisponde meglio all'opinione "media" degli altri giocatori. Il tradurre questa opinione media in numero di puntate è affare del modello.

Il criterio con cui scegliere questo pronostico, non è quindi che esso sia "buono" o "cattivo" nel senso di indovinare il risultato ma di essere più o meno aderente al comportamento degli scommettitori, che noi potremo valutare dal fatto che il nostro modello di calcolo del numero di puntate risulti più o meno buono.

Sarà necessariamente un pronostico pubblicato da qualche giornale, oppure potrebbe essere una qualche media dei pronostici pubblicati da diversi giornali ma sarebbe molto azzardato pensare di poter fare noi stessi un pronostico di questo tipo.

Potremmo avere qualche idea su come "ritoccare" un pronostico pubblicato per renderlo più rappresentativo della opinione media, ma sono indicazioni di cui non possiamo essere certi.

La cosa migliore è prenderne uno, od una media, fissi e stabili e continuare con quelli. Tra l'altro, se anche questo pronostico contiene un vizio sistematico, per esempio che il pronosticatore scelto sia, rispetto alla media dei giocatori, tendenzialmente più scettico sulla probabilità dei 2, questo vizio dovrebbe venire almeno in parte corretto dal modello stesso, che è stato determinato in base a questi dati. Quindi se anche il pronostico è "viziato", se questo "vizio" rimane stabile, tutto va bene, dovremo quindi continuare, per usare questo metodo, ad utilizzare il pronostico della "Repubblica" sul quale è stato determinato il modello.

Dobbiamo solo sperare che il pronosticatore o i pronosticatori della "Repubblica" continuino sempre a ragionare nello stesso modo di prima, che possibilmente non leggano questo libro o che comunque, se lo leggono, non si creino dei problemi e continuino tranquillamente come prima. Tra l'altro pronosticano abbastanza bene.

Il secondo uso del pronostico

Il secondo uso che noi facciamo del pronostico è per calcolare la probabilità di uscita di una determinata colonna.

Qui il discorso, almeno in linea di principio, è completamente diverso: ci serve un pronostico "tecnico" od "oggettivo" che sia il migliore possibile, cioè che corrisponda nel modo migliore alla effettiva probabilità di verificarsi dei diversi risultati.

Quando il pronostico dice: "la probabilità dei risultati della partita A-B è 50%/1 — 30%/X — 20%/2", questo può essere tecnicamente interpretato come qualcosa del genere:

— Se questa partita venisse giocata, nelle stesse condizioni, per 100 volte, vincerebbe 50 volte la squadra A, 20 volte la squadra B e 30 volte ci sarebbe un pareggio. Questa è l'interpretazione di scuola calcistica italiana.

Oppure con l'interpretazione di scuola calcistica inglese:

— Questa partita è un campione casuale di 90', di azioni da 15" ciascuna, della squadra A contro la squadra B e viceversa, ciascuna azione da una parte o dall'altra ha una certa sua probabilità di andare in goal, quando l'arbitro fischia la fine chi è avanti è avanti e chi è indietro è indietro.

  Secondo una distribuzione casuale, la probabilità che abbia fatto più goals la squadra A è del 50%, che ne abbiano fatto un ugual numero è del 30%, che ne abbia fatto di più la squadra B è del 20%.

Qualunque sia l'interpretazione, quello che ci serve è il pronostico "migliore" e non è detto che sia lo stesso che usiamo per il primo scopo. In questo libro, viene usato per entrambi gli scopi il pronostico della "Repubblica" perché dovevamo comunque usare un pronostico già pubblicato, però i due pronostici sono, nell'uso pratico, indipendenti.

Se siamo amici di qualche portiere, il secondo pronostico potremmo farlo noi stessi.

Restando comunque sul terreno della legalità e della tecnica, possiamo comunque tentare di ottenere il migliore pronostico prendendo un pronostico pubblicato, oppure il nostro stesso pronostico, e facendone una "taratura", cercando cioè di individuare e correggere, se ci sono, i difetti sistematici, oppure facendo una "valutazione" di diversi pronosticatori per scegliere il migliore.

Questi sono gli argomenti dei due prossimi capitoli.

CAPITOLO 6

LA TARATURA DEL PRONOSTICATORE

La taratura del pronosticatore consiste, come per la taratura di un qualsiasi strumento, di verificare se esiste un errore sistematico nel valore indicato rispetto alla realtà ed introdurre una correzione nella scala di lettura.

Per esempio, per tarare il tachimetro di un'automobile, possiamo cronometrare il tempo per percorrere un chilometro, o meglio ancora due chilometri, lanciati in autostrada a velocità costante e quindi controllare se il tachimetro è esatto oppure se dobbiamo aggiungere o togliere alla lettura del tachimetro qualche chilometro per avere il valore vero della velocità. Per il pronosticatore non abbiamo la possibilità di fare la taratura in questo modo, ma possiamo fare una verifica sulla successione dei suoi pronostici, lo scopo è lo stesso: verificare se non commetta un errore sistematico.

Il criterio è questo:

In primo luogo abbiamo bisogno di una serie abbastanza lunga di pronostici del pronosticatore che vogliamo tarare. Supponiamo che in questa serie, il pronosticatore abbia assegnato per 50 volte una probabilità del 10% ad un qualsiasi risultato; se questa valutazione è corretta, il risultato si verificherà nel 10% dei casi e dovremmo riscontrarlo circa 5 volte nell'intera serie. Se la frequenza con cui il risultato si è verificato è più alta, potremmo dire che il pronosticatore tende a "sottovalutare" la probabilità del terzo risultato, "sopravvalutando" in un certo senso se stesso, mentre al contrario se la frequenza è più bassa potremmo dire che tende a stare troppo sul prudente.

Globalmente, se il pronostico è corretto, dovremmo avere per tutti i valori di probabilità utilizzati, una frequenza percentuale di realizzazione uguale alla probabilità stessa.

Occorre ovviamente che il pronosticatore "ci dica qualche cosa" perché si verifica facilmente che il "pronosticatore nullo", che assegna

probabilità 33-33-33, è sempre automaticamente e perfettamente verificato.

Per una taratura attendibile è necessario un gran numero di pronostici, in effetti la taratura funziona soltanto, ed eventualmente, per i valori sufficientemente rappresentati, se in una classe ci si può mediamente aspettare una frequenza di, ad esempio, 5 risultati, non potremmo trarre conclusioni sicure dal trovarne 7 oppure 3, sono oscillazioni che possiamo tranquillamente attribuire al caso. Verifiche quantitative del tipo di quelle del test dell'ipotesi nulla, sono evidentemente possibili, ma non sono state eseguite. Dopo tutto per questo pronostico per noi non è tanto importante "conoscere" quando "decidere", e possiamo eventual-

Probabilità	Assegnata n. casi	Realizzata n. casi	Frequenza relativa
5	16	0	0
10	44	3	6
15	48	9	18
20	106	22	20
25	66	21	31
30	118	43	36
33	51	17	33
35	65	21	32
40	111	41	36
45	39	22	56
50	87	33	37
55	8	3	37
60	43	24	55
65	4	4	100
70	8	5	62
75	3	3	100
80	1	1	100
85	1	1	100
90	0	0	—
95	0	0	—

Figura 13. Taratura del pronostico (classi originali)

mente accontentarci di indicazioni anche non del tutto certe.

Abbiamo applicato questo criterio di taratura ai pronostici della "Repubblica" per i 21 concorsi presi in esame.

I risultati sono nelle figure 13 e 14. Nella figura 13 sono riportati i risultati cosi come si sono realizzati; i risultati stessi sono molto erratici: i dati sono troppo pochi per rappresentare adeguatamente tutte le classi di valori di probabilità usate dal pronosticatore.

Nella figura 14, le classi sono state raggruppate a due a due, per avere una maggiore significatività.

Da questa figura sembra comunque di poter concludere che il pronostico non presenta sostanziosi vizi. Ai nostri fini di elaborare un pronostico nostro potremmo forse spingerci un po' più in là nel caso del primo e terzo risultato; sembra che il pronosticatore potrebbe essere meno prudente, viceversa nelle gamme centrali, probabilità fra il 25 ed il 60%, quando i tre risultati hanno probabilità vicine fra loro, sembra che il pronosticatore enfatizzi le differenze al di là della loro reale consistenza. Si tratta comunque di scostamenti molto modesti.

Il programma di taratura è riportato in allegato.

Probabilità	Assegnata n. casi	Realizzata n. casi	Frequenza relativa	
5 — 10	16+ 44= 60	0+ 3= 3	5	terzo
15 — 20	48+106=154	9+22=31	20	risultato
25 — 30	66+118=184	21+43=64	35	
33	51	17	33	primo-secondo
35 — 40	65+111=176	21+41=62	35	-terzo
45 — 50	38+ 87=126	22+33=55	44	risultato
55 — 60	8+ 43= 51	3+24=27	53	
65 — 70	4+ 8= 12	4+ 5= 9	75	primo
75 — 80	3+ 1= 4	3+ 1= 4	100	risultato
85 — 90	1+ 0= 1	1+ 0= 1	100	

Figura 14. Taratura del pronostico. (classi raggruppate)

CAPITOLO 7

LA VALUTAZIONE DEL PRONOSTICATORE

Un problema diverso e più interessante è quello di valutare la qualità di un pronostico probabilistico, cioè la sua attendibilità, senza pensare a migliorarlo.

Teniamo conto che al nostro pronosticatore si richiede di assegnare una probabilità a ciascuno dei tre risultati possibili, quindi la singola partita non ci dice niente poiché in ogni caso uscirà uno dei tre risultati possibili; se il pronosticatore ha assegnato probabilità 60-30-10 e si verifica il terzo risultato, questo non significa che il pronostico fosse un cattivo pronostico, le probabilità potevano in effetti essere quelle giuste.

Certamente però, se si verificasse regolarmente il risultato a cui il pronosticatore assegna una probabilità del 10%, avremmo seri motivi di dubitare della affidabilità del suo pronostico.

Dovremo quindi stabilire una regola per "dare il voto" al pronosticatore sulla base del suo pronostico e del risultato effettivo e vedere, dopo un certo tempo, il voto medio risultante.

Prendiamo in esame la figura 15 che è simile al diagramma ternario già usato: un pronostico probabilistico è rappresentato da un punto all'interno del triangolo equilatero (lati e vertici compresi). La probabilità di ciascuno dei tre risultati è rappresentata dalla distanza del punto-pronostico dalla base opposta al vertice corrispondente al risultato in questione.

Se prendiamo come unità di misura l'altezza del triangolo, la somma delle tre distanze è sempre uguale ad uno.

Nella figura 15, il punto pronostico corrisponde alle probabilità P1, PX, e P2 assegnate rispettivamente ai risultati 1, X e 2.

Il voto che assegneremo sarà uguale al quadrato della distanza del punto-pronostico dal vertice corrispondente al risultato che si è effettivamente verificato. Questo metodo di valutazione è la "regola di Brier";
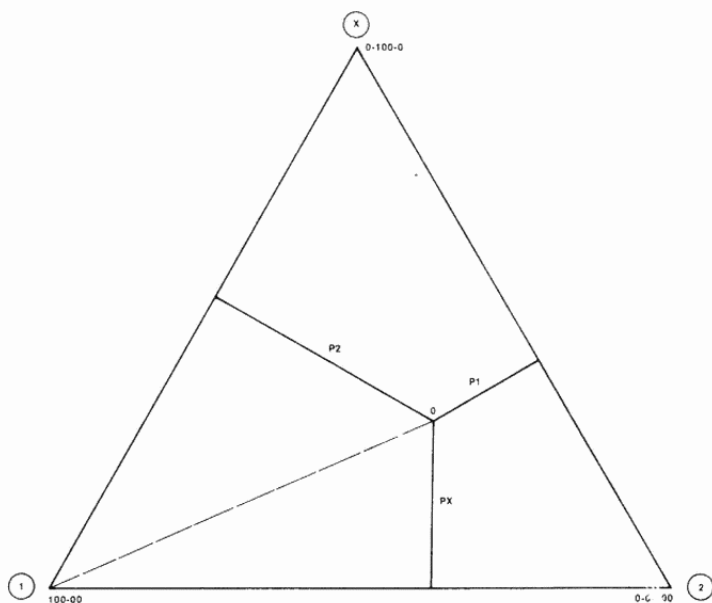
Figura 15 — Rappresentazione del pronostico probabilistico.

il voto assegnato ha il significato di un errore, nel senso che, migliore è il pronostico più basso è il voto.

Nella figura 15, supponendo che si sia verificato il risultato 1, il segmento tratteggiato rappresenta la distanza fra il punto-risultato (1) ed il punto-pronostico (0), il quadrato di questa distanza sarà il voto assegnato al pronosticatore per questo pronostico.

Analiticamente, la formula è inserita nel programma allegato, il voto si calcola con:

– Se si verifica il risultato 1 :
S1=PX↑2+(PX*Tg(30)+P2/Cos(30))↑2
– Se si verifica il risultato X :
SX=P1↑2+(P1*Tg(30)+P2/Cos(30))↑2
– Se si verifica il risultato 2 :
S2=P1↑2+(P1*Tg(30)+PX/Cos(30))↑2

È bene dire subito che questo procedimento permette di confrontare fra loro diversi pronosticatori che esercitino i loro pronostici sullo stesso processo,quindi ci potrà servire per scegliere il nostro pronosticatore prediletto oppure per confrontare la nostra abilità con quella di altri pronosticatori.

Non ci permette invece di dare una valutazione assoluta di un singolo pronosticatore, se non conosciamo le caratteristiche del processo al quale applica il suo pronostico.

Possiamo però stabilire almeno alcuni valori limite, riferendoci ad un singolo pronostico.

Un primo valore limite è il pronostico "perfetto": se il pronosticatore ha assegnato probabilità 100-0-0, e si verifica effettivamente il risultato al quale è stata assegnata probabilità uguale a 100 il voto-errore per questo pronostico è uguale a zero.

Un altro valore limite è il pronostico "completamente errato": se il pronosticatore ha assegnato probabilità 100-0-0 e si verifica uno dei due risultati ai quali è stata assegnata una probabilità nulla, il voto-errore assegnato sarà 1.333333.

Un terzo caso limite interessante è quello del "pronosticatore nullo", o pronostico a informazione nulla, che ci fornisce una probabilità di 33-33-33 che è come dire "non sono in grado di dire nulla".

Per questo pronostico, qualunque sia il risultato, avremmo un voto-errore di 0.4444444, essendo S1=SX=S2=0.444444444.

Passando a considerare il risultato medio atteso dopo un certo numero di prove, se sono F1, FX ed F2 le frequenze con cui si sono verificati i risultati 1, X e 2, il punteggio finale medio è dato da:

F1*S1+F2*S2+FX*SX

È facile vedere che per il "pronosticatore nullo", poiché S1=SX=S2=0.44444 ed F1,FX,F2 hanno sempre come somma 1, il voto sarà sempre 0.4444444 per qualunque valore particolare di F1, FX ed F2.

Quindi, il pronosticatore ad informazione nulla, qualunque sia il processo, riceve sempre un voto di 0.44444444.

Questo valore si può quindi considerare un valore "massimo", in quanto, di fronte ad un qualsiasi processo, il pronosticatore può sempre confessare la sua impotenza assegnando probabilità 33-33-33 ed ottenere quindi il punteggio di 0.44444.

Non che non si possa fare di peggio, a questo punto dovremmo parlare di un pronosticatore "negativo"; ad esempio se il processo è effettivamente del tutto imprevedibile con probabilità dei tre risultati 33-33-33, che potremmo chiamare per analogia un processo "ad informazione nulla", se il pronosticatore assegna probabilità diverse da 33-33-33 il suo punteggio sarà più alto. Una delle proprietà di questo metodo di valutazione è infatti che, per una data terna "di processo" F1, FX, F2, il punteggio minimo si ottiene se P1=F1, PX=FX e P2=F2, cioé se il pronosticatore assegna le vere probabilità del processo.

Peggio di 0.44444, può anche ottenere il pronosticatore "millantatore": ad esempio, un pronosticatore che non abbia alcun elemento di giudizio e, anzichè confessarlo con un onesto 33-33-33, mascheri la sua ignoranza assegnando a caso con uguali frequenze una probabilità del 100% ad uno dei tre risultati e zero agli altri due, otterrà nel lungo andare un voto-errore di 0.8888888, anche in questo caso qualunque sia il processo.

Questi casi limite numerici, ci danno dei larghi criteri di riferimento. Ma passiamo ora a considerare delle situazioni più realistiche.

I voti-errore che ci interessano stanno fra 0 e 0.444444 che è il voto del pronosticatore "nullo".

Abbiamo visto prima, che il voto migliore che un pronosticatore possa ottenere si realizza quando vengono assegnati valori di probabilità ai tre risultati, uguali alle probabilità effettive che ha il processo di generare questi risultati.

Il voto ottenibile non dipende però soltanto dalle sue abilità ma anche dalla effettiva prevedibilità del risultato, cioé dalle caratteristiche del

processo che sono espresse dalle probabilità F1, FX ed F2. Se il processo è totalmente imprevedibile, cioé se queste probabilità sono 33-33-33, il miglior risultato che può ottenere il pronosticatore è di 0.4444444 e si ottiene assegnando effettivamente probabilità 33-33-33, saremmo in presenza di un processo ad "informazione nulla" e ad un pronosticatore "perfetto".

Tuttavia, se il processo ha una sua informazione, cioé se oggettivamente i tre risultati non sono ugualmente probabili, il pronosticatore può ottenere risultati medi migliori.

Ad esempio, se ogni volta uno dei tre risultati avesse 100% di probabilità di verificarsi, ed il pronosticatore assegnasse 100% di probabilità a questo risultato, il suo errore finale sarebbe 0; saremmo in presenza di un processo "perfettamente prevedibile" o a "informazione perfetta" e ad un pronosticatore ugualmente "perfetto".

In pratica, noi non conosciamo il livello di "prevedibilità" del campionato di calcio, ma sappiamo che il punteggio del pronosticatore non può essere migliore di quello che il processo stesso gli consente, pertanto un determinato voto-errore conseguito da un pronosticatore può significare o un pronostico "perfetto" su un processo che consente quel valore massimo oppure un pronostico meno perfetto su un processo che consentirebbe in realtà un errore più basso.

Nella figura 16, sono riportati i valori massimi conseguibili per diversi livelli di "prevedibilità" del processo. Il diagramma riporta l'intero triangolo per maggiore evidenza grafica, benché i tre risultati siano ovviamente simmetrici. I vertici rappresentano un ipotetico campionato in cui uno dei tre risultati sia sistematicamente prevedibile con probabilità del 100%; in queste condizioni l'errore minimo è 0.

Il baricentro rappresenta invece una condizione di totale imprevedibilità nella quale il voto minimo ottenibile è come già visto 0.4444444.

Gli altri valori corrispondono a valori intermedi di probabilità assegnabili al risultato più probabile ed altri due.

I punti mediani di ogni lato rappresentano condizioni di prevedibilità in cui fosse possibile scartare uno dei tre risultati, assegnandogli probabilità uguale a 0, ed assegnare 50% ciascuno agli altri due.

In queste condizioni, il voto minimo ottenibile è di 0.3333333.

Ci sembra di poter dire che questa condizione corrisponde ad una
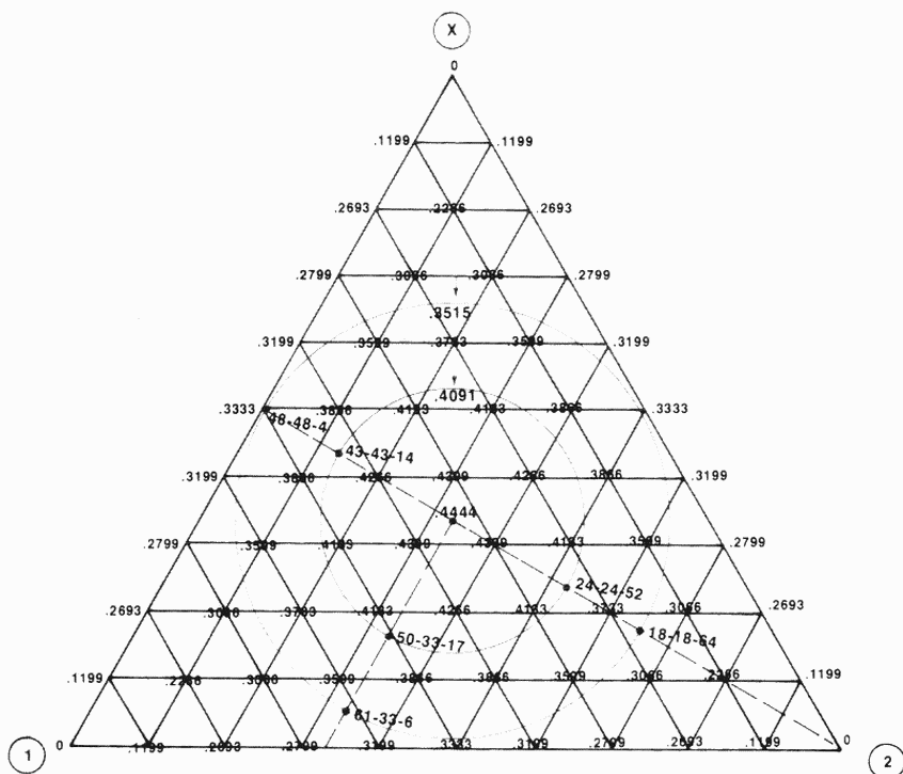
Figura 16 — Errori minimi ai diversi livelli di prevedibilità.

"prevedibilità" ben superiore a quella consentita dal campionato di calcio. In effetti, se questa condizione fosse vera, sarebbe possibile con "tredici doppie" avere la certezza di vincere sempre.

Le considerazioni che abbiamo esposto servono, come già detto, a confrontare fra loro pronosticatori diversi oppure a confrontare le prestazioni nel tempo di uno stesso pronosticatore, non possiamo dare una valutazione assoluta, in quanto non conosciamo il livello di informazione del "processo - campionato".

I tests sono però una grande tentazione, almeno come gioco di società. Ancora meglio se si possono fare a qualcuno che non lo sa, e poi spararagli davanti i risultati.

Abbiamo quindi sottoposto al test l'ignoto pronosticatore della "Repubblica" che era l'unico di cui disponessimo, con un numero di pronostici sufficiente ad evitare grossi errori casuali.

Ci scusiamo dicendo che abbiamo anche ritenuto interessante poter disporre di un primo dato che possa servire di confronto per altre valutazioni.

Abbiamo utilizzato i soliti 21 concorsi ed abbiamo il piacere di registrare che, dopo un comprensibile inizio incerto che lo ha portato a sfiorare il campo dei pronosticatori "negativi", il nostro uomo ha migliorato e concluso con un voto medio generale di 0.4091.

Come prestazione di punta, su tre concorsi consecutivi, abbiamo registrato nei concorsi 27-30 l'entusiasmante media di 0.3515, un passo sicuramente insostenibile a lungo, l'atleta ha infatti pagato lo sforzo con un cedimento nel finale.

Nella figura 16, sono state tracciate le circonferenze relative ai livelli di errore riscontrati. Il livello di 0.4091 è probabilmente migliorabile, però chi pensasse che questi risultati siano scarsi, può esaminare la figura 16 e considerare le intersezioni di queste circonferenze con una qualunque delle altezze del triangolo e con una parallela ad una base passante per il baricentro (linee tratteggiate in figura).

Tutti i punti che stanno su una circonferenza rappresentano configurazioni di uguale "informazione del processo" cioè di uguale "prevedibilità".

Le due intersezioni con l'altezza rappresentano rispettivamente i casi, di più facile visualizzazione, in cui uno dei tre risultati ha la probabilità massima, e rispettivamente minima, e la probabilità residua si ripartisce ugualmente fra gli altri due risultati, mentre l'intersezione con la parallela alla base rappresenta il caso in cui il secondo risultato ha la probabilità del 33% ed il primo ed il terzo sono ai valori corrispondenti al livello di informazione del processo, espresso dal diametro del cerchio.

Il cerchio 0.4091 rappresenterebbe quindi una prevedibilità "media generale" del campionato corrispondente a:

- un risultato più probabile, gli altri due equivalenti:
  52% – 24% – 24%
- un risultato meno probabile, gli altri due equivalenti:
  43% – 43% – 14%
- probabilità decrescenti dal più al meno probabile:
  50% – 33% – 17%

Il cerchio a 0.3515 corrisponderebbe invece ai valori:

- un risultato più probabile, gli altri due equivalenti:
  64% – 18% – 18%
- un risultato meno probabile, gli altri due equivalenti:
  49% – 48% – 4%
- probabilità descrescenti dal più al meno probabile:
  61% – 33% – 6%

Se teniamo conto che si tratta di prevedibilità "media", su tutte le partite di tutto il campionato, dalle prime alle ultime giornate, anche i valori al livello 0.4091 non sembrano, almeno a sensazione, tanto male.

Quanto ai valori al 0.3515, sempre a sensazione, sembrano corrispondere ad un livello di informazione media che il campionato di calcio è ben lontano dall'avere e dobbiamo prudentemente attribuirli al caso.

Qualunque record ha comunque delle componenti di casualità.

PARTE TERZA

IL PROGRAMMA TOTOMAC IN ESERCIZIO

CAPITOLO 8

IL PROGRAMMA TOTOMAC ESECUTIVO

Il programma Totomac esecutivo si compone di diversi moduli:

MODULO 1. Riceve in Input il pronostico "Repubblica" della settimana in questione e memorizza la tabella delle probabilità limite.
La procedura è quella indicata nella parte prima. È raccomandato l'uso del pronostico "Repubblica", in quanto il modello è tarato su questo.

MODULO 2. Riceve in input un eventuale pronostico "personale" (vedi parte seconda).
Esegue a ripetizione:
SOTTOMODULO 2.1.
Genera una colonna, fino ad un numero totale di colonne preordinato.
SOTTOMODULO 2.2.
Verifica se la colonna è "buona", la memorizza o la scarta, e ritorna al Sottomodulo 2.1.

MODULO 3. Stampa le colonne prodotte ed accettate.

Il modulo 2. si compone quindi di due sottomoduli, relativi rispettivamente alla generazione delle colonne ed alla loro verifica ed accettazione.

La verifica ed accettazione operano sostanzialmente secondo la regola base:

- Viene calcolata la probabilità (P1) di uscita della colonna generata dal sottomodulo 2.1, secondo il pronostico "personale".
- Viene calcolata la frequenza di giocata (F1) secondo il modello incorporato nel sistema ed il pronostico "Repubblica".
- La colonna viene accettata se il rapporto tra P1 ed F1 è favorevole, cioè maggiore di uno, e respinta in caso contrario.

Il sottomodulo 2.2. prevede però altri parametri di controllo che saranno esaminati al capitolo seguente.

In questo capitolo discuteremo invece brevemente il sottomodulo 2.1.

E cioé la metodologia seguita nella generazione delle colonne.

Un primo metodo possibile consiste nel generare queste colonne completamente "a caso", spetterà interamente al sottomodulo 2.2. scartare le colonne non soddisfacenti.

Questo è il più antico e collaudato metodo che si conosca per ottenere un prodotto stabile: "produrre a caso e ammazzare il fuori-standard"; le gazzelle lente, le mangiano i leoni e rimangono solo quelle veloci.

È vero che il risultato che abbiamo tutti sotto gli occhi può non sembrare, a volte, perfettamente riuscito, però questo sarebbe anche il metodo migliore: tutto l'onere della selezione è lasciato al modulo 2.2 e non esiste nessun filtro preventivo che limiti le possibilità di realizzazione.

Dal nostro punto di vista, questo metodo ha però l'inconveniente della lentezza, noi non abbiamo i millenni a disposizione e nemmeno le ore.

In un concorso normale, le colonne che passeranno l'esame del modulo 2.2. stanno in gran parte nelle 300-400.000 più probabili, cioè nel primo 20-25%. Possiamo rivedere la figura 11 e rileggerci le osservazioni conclusive sul modello. La generazione "a caso", che produce con ugual probabilità qualunque tipo di colonna, produrrebbe quindi mediamente il 75-80% di colonne sistematicamente non interessanti e quindi richiederebbe un tempo eccessivo per la generazione ed accettazione del numero di colonne voluto.

È quindi opportuno ricorrere ad un altro metodo di generazione. Il metodo che abbiamo adottato è un compromesso che ci sembra soddi-

sfacente.

L'operatore introduce, per ciascuna partita, i risultati 1,X,2 che vuole siano presi in considerazione, secondo lo schema abituale di definire una partita come una fissa, una doppia od una tripla.

Il sistema genera tutte e sole le colonne che corrispondono allo schema e le passa all'esame del modulo 2.2; naturalmente non tutte le colonne verranno accettate. Finita l'esplorazione delle colonne richieste dall'operatore, avremo un certo numero di colonne accettate, se questo numero è scarso il processo potrà essere ripetuto con uno schema iniziale più allargato, se è eccessivo può essere ripetuto con uno schema più ristretto.

Il numero di colonne che verranno generate, e quindi il tempo che sarà necessario, è facile da calcolare: è il prodotto del numero di risultati previsti per ciascuna partita. Ad esempio, se viene inserito uno schema con 6 triple e 4 doppie, verranno generate

$$3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 2 \times 2 \times 2 \times 2 = 11664 \text{ colonne,}$$

che è un numero rilevante ma non proibitivo.

Naturalmente non tutte queste colonne verranno accettate dal modulo 2.2., è impossibile dire in generale quale sarà la frazione accettata in quanto essa dipende dalla struttura del concorso e dalla "griglia" di selezione (vedi capitolo seguente) che verrà imposta al modulo 2.2.

È chiaro che questo procedimento non è l'ottimo, in quanto l'operatore, limitando inizialmente le colonne da considerare, ne esclude una gran parte che sarebbe stato meglio far controllare al modulo 2.2.; tuttavia non tutto si può avere.

È comunque una buona regola, compatibilmente con il tempo disponibile e con la velocità del mezzo di calcolo, stare quanto più larghi si può nella generazione di colonne e più stretti nella griglia di accettazione del modulo 2.2.

La colonna rappresentata come un numero in base tre

Nel programma di generazione, una colonna è rappresentata da un numero; in questo modo i tredici simboli che compaiono in una colonna possono essere contenuti in uno spazio di memoria inferiore.

Questo viene effettuato considerando una colonna come un numero in una numerazione in base tre.

Questa nostra numerazione dispone di tre simboli 1, X e 2, che esprimono quantità, rispettivamente, 0, 1 e 2.

La colonna, scritta in orizzontale con la tredicesima partita all'estremo di sinistra e la prima all'estrema destra, rappresenta un numero, in questa numerazione, ed ogni diversa colonna rappresenterà un diverso numero.

La posizione di un simbolo determina la quantità che esso rappresenta.

Nella consueta numerazione decimale, una cifra posta in una determinata posizione si intende moltiplicata per 10 elevato ad un esponente uguale alla sua posizione contata dalla destra; la prima posizione a destra ha valore zero.

Nella nostra numerazione ternaria, ogni simbolo si intende analogamente moltiplicato per 3 elevato all'esponente uguale alla sua posizione, sempre contata da destra.

In questo modo ad esempio, una colonna di tutti 1 avrà valore 0, una colonna con tutti uno ed un solo X nella terza posizione avrebbe il valore decimale 9, eccetera, una colonna con tutti 2 corrisponde al valore decimale 1.594.322 cioè 3↑13–1.

CAPITOLO 9

IL CONTROLLO DEI PARAMETRI DEL SISTEMA TOTOMAC

Una volta generate le colonne, il sottomodulo 2.2. ha il compito di separare le colonne "buone" dalle colonne "cattive", per accettare le prime e respingere le seconde.

Il criterio base del Totomac consiste, come sappiamo, nel rapporto tra la probabilità di uscita di una colonna e la frequenza prevista di puntate su quella stessa colonna: più alto è questo rapporto, migliore è la colonna.

Nell'ambito di questo criterio base, c'è tuttavia ancora spazio per diverse strategie che ciascuno può definire secondo le proprie preferenze personali.

In generale, la strategia personale opererà su due parametri:

– il "valore atteso"

– La "probabilità di rovina"

Intuitivamente nel senso di rendere massimo il primo e minima la seconda, questo però è possibile solo nei limiti che il Totocalcio ci mette a disposizione, quindi si tratterà di realizzare un compromesso secondo i nostri gusti e preferenze che si sintetizzano in quella che viene chiamata "propensione al rischio".

Il Totomac ci permette di controllare separatamente e indipendentemente questi due parametri in modo da ottenere la strategia che ci è più congeniale: vediamo come operano questi due parametri ed i loro effetti.

Il "valore atteso"

Il "valore atteso" di una colonna, di cui abbiamo già parlato, è dato dalla probabilità che esca la colonna giocata moltiplicato il premio che eventualmente si vince; siccome la eventuale vincita è il monte premi diviso il numero di giocate raccolte dalla colonna, il "valore atteso" è controllato direttamente dal rapporto tra probabilità di uscita e frequenza di giocata.

Per chiarire il concetto, se ce n'è bisogno, aggiungiamo che un "valore atteso" alto non significa necessariamente una vincita alta in valore assoluto, 100.000.000 di lire piuttosto che 100.000, ma soltanto che la colonna è "poco giocata" rispetto alla sua probabilità. Può essere una colonna molto probabile, che sarà quindi anche molto giocata, ma sempre meno della sua probabilità, oppure una colonna poco probabile ed ancor meno giocata.

Comunque il "valore atteso" determina il risultato economico che noi ci possiamo aspettare dal nostro gioco nel lungo termine.

Il modulo 2.2. ci permette di scegliere il "valore atteso" minimo cioè il valore limite del rapporto tra probabilità e frequenza di giocata, al di sotto del quale le colonne vengono respinte ed al di sopra del quale vengono accettate.

È evidente che noi possiamo scegliere un valore altissimo di questo rapporto, supponiamo colonne che siano almeno dieci volte più probabili che giocate, ma non possiamo creare queste colonne se non esistono.

Se scegliamo un valore limite molto alto, saranno molto poche le colonne accettate, quindi noi avremo bensì scelto "il-meglio-del-meglio", ma le colonne accettate saranno molto poche, avremo un "valore atteso" alto ma una probabilità di vincita globalmente bassa.

I valori accettabili dipendono dallo specifico concorso, in alcuni concorsi possono esserci delle prospettive migliori, in altri meno buone, in generale valori limite del rapporto compresi fra 1.1 e 2.0 sembrano ragionevoli.

In ogni modo è sempre possibile fare un primo giro di Totomac per un concorso specifico ed eventualmente ripeterlo con valori limite diversi se il numero di colonne accettato non è soddisfacente.

La "probabilità di rovina"

Il secondo parametro concerne la "probabilità di rovina". Nonostante il nome catastrofico questo è un concetto tecnico.

Si intende per "rovina" l'esaurimento delle risorse complessive stanziate da un giocatore dopo un certo numero di giocate senza vincita o con vincite insufficienti.

Praticamente questo al Totocalcio significa piuttosto se vogliamo tenerci a risultati più probabili limitando le puntate, oppure se vogliamo spingerci anche a quelli meno probabili puntando di più.

Questo viene quindi controllato dalla sola variabile "probabilità di uscita della colonna".

Il modulo 2.2. presuppone che nessuno voglia escludere le puntate più probabili e concentrarsi su quelle meno probabili, per quanto pingui possano essere: vincere al Totocalcio è già abbastanza difficile per volercelo rendere ancora più difficile.

Si richiede quindi all'operatore di indicare una probabilità "minima" al di sotto della quale le colonne vengono comunque scartate, anche se il loro "valore atteso" fosse buono. Anche qui i valori accettabili dipendono dallo specifico concorso, e non possiamo creare colonne ad alta probabilità se non esistono, se vogliamo giocare molte colonne è giocoforza porre un limite più alto.

In generale possiamo dire che probabilità di 20-30 per milione sono molto alte, rischiamo di accettare pochissime colonne; 1 per milione è ragionevolmente basso. Si può notare dalla figura 11 che in nessuno dei 21 concorsi presi in esame si è avuta una colonna "buona" con probabilità inferiore a 0.5 per milione circa, corrispondente a 75 vincitori "equi", quindi non sembrerebbe raccomandabile scendere al di sotto; tuttavia niente ci permette di escludere che questa situazione si verifichi in futuro.

La "riduzione" del sistema

Una terza possibilità offerta dal Totomac è quella di operare la "riduzione" del sistema giocato: questo consiste nell'eliminare le colonne che presentano un solo segno diverso da una colonna già accettata.

Questa procedura è ben nota e serve ad evitare una eccessiva concentrazione delle giocate che, in caso favorevole, darebbe vincite molto forti (13 e 12), ma che per contro riduce il numero di casi favorevoli.

Questa scelta è anch'essa rilevante dal punto di vista della "probabilità di rovina" e non da quello del "valore atteso".

CAPITOLO 10

COMMENTI ALLE SOLUZIONI ADOTTATE

Questo capitolo contiene alcune considerazioni per un eventuale miglioramento del sistema Totomac.

Evidentemente si può pensare a delle modifiche operative:

- Memorizzazione delle colonne giocate per poter effettuare col calcolatore annche il controllo della vincita.
- Stampa diretta delle colonne prodotte sulla schedina.
- Diversa concezione del "generatore" di colonne (modulo 2.1.)
- Diversa gestione del "filtro" di verifica (modulo 2.2.)

Tuttavia non è di queste che intendiamo parlare, quanto del "modello" che calcola il numero di puntate, che costituisce l'elemento tipico del Totomac.

Il modello può essere migliorato per ottenere una maggiore precisione nel calcolo: si tratta di spingere ulteriormente l'elaborazione dei dati o di raccoglierne altri e considerare anche quelli. Ora come ora, il modello discrimina abbastanza bene fra colonne "sopra" e "sotto-puntate", ma potrebbe essere più preciso nel calcolo dell'effettivo numero di puntate.

Questo permetterebbe un calcolo del valore atteso che, allo stato attuale del modello, non sarebbe molto attendibile.

Siamo convinti che si possa arrivare ad una precisione del tutto soddisfacente del modello anche sotto questo punto di vista, naturalmente questa è solo una supposizione, ma 150 milioni di colonne alla settimana sono un campione talmente grosso da poter sperare di lasciare a variazioni casuali un margine molto più ristretto di quello attuale.

Un altro motivo per giustificare un ulteriore studio del modello è che il

comportamento degli scommettitori cambierà nel tempo, non rapidamente perché 150 milioni di colonne giocate sono una massa difficile da muovere, ma nel corso degli anni certamente cambierà.

Riprendiamo quindi in questo capitolo finale alcune osservazioni sparse qua e là nel testo, ed in particolare nel capitolo introduttivo della prima parte, su quelle che ci sembrano le linee più promettenti per rendere il modello più esauriente e più preciso.

1)

Senza introdurre altre variabili nel modello e conservando la stessa struttura alla formula di calcolo, i coefficenti costanti che intervengono nella formula stessa possono essere calcolati in modo più preciso: la determinazione attuale è stata fatta praticamente ad occhio e senza molti tentativi.

Non ci sarà molto da guadagnare, perché la valutazione ad occhio non sbaglia in genere granché, ma qualche cosa sicuramente si può ottenere per questa via.

2)

Introducendo altre variabili nel modello, ci sono degli interessantissimi candidati, quasi sicuramente rilevanti. Anche qui si tratta di supposizioni che andranno verificate nei fatti, ma un po' di ragionamento preliminare sui meccanismi non dovrebbe dare indicazioni del tutto fuori posto.

Una prima variabile è il "grado di concentrazione" o, se vogliamo, il grado di "piattezza" del pronostico.

Si è tentato di inserirlo nel modello in qualche forma, per poi rinunciare di fronte ai risultati già acquisiti ed alla scarsezza dei dati disponibili, ma quasi sicuramente avrebbe da dire la sua.

Per spiegarci meglio: il modello Totomac dice: "Tutti i giocatori giocano a partire dalla colonna più probabile e si spingono più o meno avanti verso le meno probabili, secondo la somma stanziata". Tuttavia possiamo chiederci: esiste veramente questa "colonna più probabile", sulla quale tutti i giocatori sono d'accordo?

Naturalmente questo va inteso in senso statistico: la "colonna più

probabile'' può essere in realtà una decina di colonne simili, la situazione non cambia molto. Ma possiamo aggiungere che in un concorso "ripido", in cui ogni partita abbia uno dei risultati prevalente rispetto agli altri, esiste verosimilmente questa "colonna più probabile" e con buona approssimazione tutti o quasi gli scommettitori partono da quella. Viceversa, in un concorso "piatto", al limite in un concorso nel quale tutte le partite abbiano come pronostico 33-33-33, cioè il massimo della "piattezza' o il minimo della "informazione", la colonna più probabile non esiste affatto: hanno tutte la stessa probabilità.

È quindi prevedibile, che in un concorso "piuttosto-piatto" non ci sia quell'accordo generale su quale sia la colonna più probabile, i giocatori partiranno un po' da una colonna ed un po' da un'altra.

La "piattezza" del concorso può essere misurata dal numero di colonne, a partire dalla più probabile, necessarie a sommare il 20% della probabilità totale (il metodo Montecarlo ci costruisce l'integrale della probabilità e serve egregiamente a calcolare questo parametro indice). In condizioni di concorso "piatto", a rigor di logica, dovrebbe acquistare peso nel modello la variabile "probabilità della colonna", a fianco della variabile "posizione della colonna nella serie", e soppiantarla completamente per un concorso totalmente "piatto".

Questa sembra la variabile più importante fra quelle trascurate e non ci stupiremmo di trovarla in una seconda release del modello.

3)

Esistono diverse categorie di scommettitori. Questa era l'ipotesi iniziale del Totomac che successivamente è stata tolta dal modello, ma il "fatto fisico" esiste sicuramente e dandone una adeguata rappresentazione il modello dovrebbe riuscire più preciso.

Attualmente la componente dei "trottolisti" interviene come correttivo della strategia dei sistemisti, ma una frazione di "trottolisti puri" che seguono la trottola senza guardare le partite meriterebbe una voce autonoma.

Altrettanto si può dire dei "sistemisti corti", che giocano poche colonne ragionate e che dovrebbero, essendo tanti, dare luogo ad uno schema legato, anche questo, più alla probabilità della colonna che alla sua posizione nella serie.

Anche se la probabilità della colonna, come prima variabile, si è dimostrata meno efficiente della posizione nella serie, può però servire a spiegare in parte la varianza residua.

Ovviamente, passare a considerare la popolazione totale di scommettitori come una somma di gruppi diversi, comporta passare ad un modello polinomiale, il ché aumenta i gradi di libertà del modello con i rischi che abbiamo già visto, tanto più se consideriamo questi gruppi percentualmente variabili nell'arco del campionato, come verosimilmente sono, ma con le dovute cautele, questa ipotesi merita un approfondimento.

4)

Aggiungiamo un'ultima considerazione, anche se suona più come una giustificazione che come una proposta.

In tutto il libro si è parlato di vincite, senza precisare se si tratti di un 13 o di un 12.

Questa non è una svista, è stata una scelta: tutta la trattazione è impostata sui 13 in quanto trattare i 12 avrebbe comportato notevoli complicazioni, ma la validità del metodo rimane, il 12 non è altro che un 13 mancato.

Abbiamo indicato, per il calcolo del valore atteso, l'espressione

P1 * MP / F1 = VA

In cui sono: P1 = probabilità della colonna; MP = Monte premi; F1 = frequenza di giocata della colonna = numero di vincitori

Le cose non cambiano se, per considerare separatamente i 12, consideriamo le due metà del monte premi, il valore atteso sarà composto di due termini:

VA = P1 * (MP/2) / F1 + P2 * (MP/2) / F2

In cui P2 è la probabilità che esca un 12 anziché un 13 ed F2 è il numero di vincitori col 12. Siccome P2 ed F2 sono, in prima approssima-

zione, uguali a P1 ed F1 moltiplicati entrambi per 26, il risultato non cambia.

Volendo essere più precisi, una differenza c'è.

Per quanto riguarda i 13, noi dividiamo la vincita con quelli che hanno giocato la nostra stessa colonna, di cui abbiamo calcolato la frequenza di giocata. Per quanto riguarda i 12 invece noi divideremo la vincita con giocatori che hanno giocato altre colonne, delle quali non abbiamo calcolato la frequenza di giocata e nemmeno possiamo pensare di farlo: possiamo fare 12 in 26 modi diversi e per ciascuno di questi ci sono altre 26 colonne, anche se non tutte diverse, che lo possono produrre.
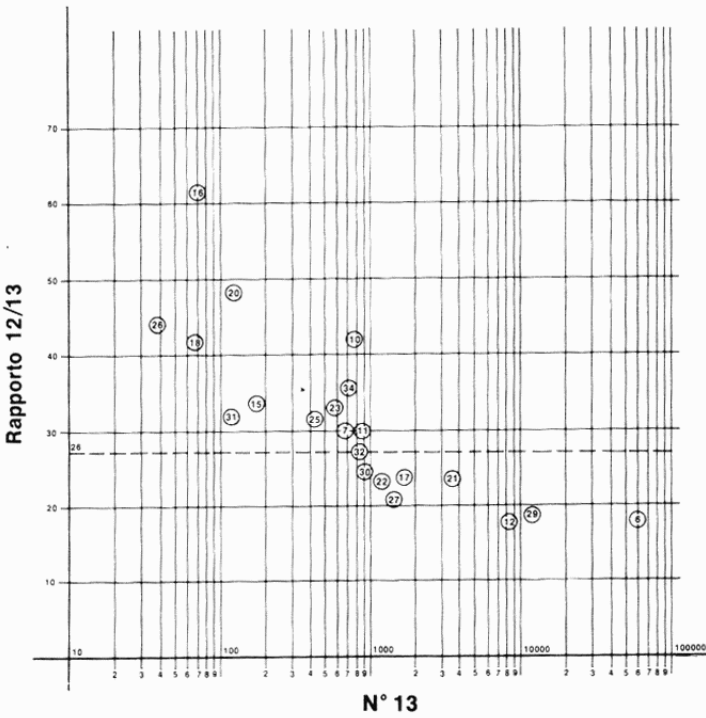
Figura 17 – Andamento dei dodici.

È però possibile una trattazione più sintetica, che si tradurrebbe in una nuova componente da introdurre nel modello, sempre in vista di raffinamenti successivi.

Esaminiamo la figura 17, questa riporta, per i nostri soliti 21 concorsi, il numero di 13 realizzati sulla colonna vincente ed il rapporto fra numero dei 12 e numero dei 13; il numero di 12 e 13 è riportato fra parentesi accanto al punto rappresentativo del concorso.

Si vede ad occhio che ad un basso numero di 13 corrisponde un alto valore del rapporto fra 12 e 13 realizzati. Questo fatto sembra avere una sua logica spiegazione: un 12 "proviene" da una colonna che aveva una sua probabilità ed una sua frequenza di giocata che potevano essere più alte o più basse della colonna effettivamente uscita, se la colonna uscita aveva una bassa frequenza di giocata, la provenienza dei 12 avviene tendenzialmente da colonne a frequenza di giocata più alta ed il rapporto fra 12 e 13 tende ad aumentare; il contrario succede se la colonna uscita aveva una frequenza di giocata relativamente alta, la provenienza avviene da colonne meno giocate e quindi il rapporto tende ad essere più basso.

Nella formula potrebbe quindi essere inserita una voce che premi le colonne a più alta probabilità in quanto il valore atteso per queste colonne dovrebbe avere una componente di provenienza 12 più sostanziosa.

CONSIDERAZIONI FINALI

"Qualunque testo di statistica o di teoria della probabilità va bene" – sentenziò Van Random – "italiano o straniero. In tutti c'è tutto quello che serve ed anche di più. E così anche sulle sciocchezze che avete infilato nel libro uno si chiarisce le idee".

"L'unico guaio, con quei testi" – proseguì – "è quella insopportabile monotonia tematica. Se alla prima tabella del libro trovi il peso alla nascita dei maialini dello Iowa, ti tocca seguire la loro vita fino alla fine del libro. In tutti i suoi aspetti. Dai più bassi ai più sublimi. Lo stesso se la prima tabella parla di sfere per cuscinetti. Sempre cuscinetti. Al massimo qualche ingranaggio."

"Meglio Fisher, il grande R.A." – continuò – "la 'progettazione degli esperimenti', o come altro si chiama. Avrei sempre voluto vedere quelle miss un po' stagionate a bere 50 tazze di tè una dopo l'altra, 25 col latte versato prima del tè e 25 col latte versato dopo e sentire se avevano ancora fiato per parlare". "Meraviglioso!" – concluse.

"E poi il Candido di Voltaire" – proseguì – "Quello non lo cita mai nessuno e intanto il naso di Pangloss è ancora tra noi. Anche chi non ti aspetteresti mai". Ci guardò un po' trucemente, la tigre che sonnecchia in lui si stava evidentemente affilando le unghie.

Eravamo alla serata conclusiva, che avevamo stabilito di dedicare alla bibliografia, ma pareva che nessuno avesse voglia di applicarsi seriamente. Riuscii comunque a indurli a sfogliare le pagine per rinfrescarsi la memoria.

"La probabilità di De Finetti nell'enciclopedia Einaudi" – intervenne Totomac che sembrava l'unico in grado di dare un contributo fattivo – "senza quell'articolo, la seconda parte non la scrivevamo mai".

"Poi W. Feller" – proseguì Totomac – "la teoria della probabilità, e la teoria dei giochi, Von Neumann e Morgenstern, Montecarlo, e Popper".

"Quello poi lo citano sempre tutti" – sbottò Van Random – "a proposito e a sproposito, e non che non ne valga la pena. Ma perché non Fisher? C'è già tutto, e con tanto di numerini per farsi i conti".

"Per favore, passami il ghiaccio" – disse Baccarat – "e non vi dispiacerebbe se cambiassimo argomento una volta per tutte?".

Tabella 1. Dati utilizzati

CONCORSO N° 6				CONCORSO N° 7			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
30	40	30	1	30	60	10	1
20	60	20	X	33	33	33	2
50	40	10	1	60	30	10	1
50	30	20	X	20	70	10	1
40	20	40	X	30	60	10	X
70	25	5	1	40	40	20	2
70	20	10	1	40	20	40	1
60	30	10	1	50	30	20	X
50	35	15	1	20	50	30	2
40	50	10	1	50	40	10	1
40	20	40	X	20	60	20	1
30	30	40	X	60	20	20	X
50	25	25	X	40	20	40	X
Giocate: 119.666.008 Vincitori: 15.922 Probab. col. vincente 0.0000047628000 N° vinc. "Equi": 569				Giocate 133.282.314 Vincitori: 66 Probab. col. vincente 0.0000002052864 N° vinc. "Equi": 27			

Tabella 1. Dati utilizzati

CONCORSO N° 10				CONCORSO N° 11			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
25	60	15	2	30	40	30	1
60	35	5	X	40	40	20	X
40	45	15	2	30	30	40	1
33	33	33	2	50	30	20	2
30	60	10	1	40	40	20	1
50	40	10	1	30	40	30	X
33	33	33	1	60	20	20	1
60	30	10	X	40	30	30	1
50	40	10	X	50	30	20	X
45	35	20	1	20	50	30	2
55	35	10	X	50	30	20	X
40	40	20	1	30	30	40	1
50	35	15	X	50	30	20	1
Giocate: 139.989.187 Vincitori: 74 Probab. col. vincente 0.0000003403764 N° vinc. "Equi": 47				Giocate: 148.910.452 Vincitori: 78 Probab. col. vincente 0.0000011197440 N° vinc. "Equi": 166			

Tabella 1. Dati utilizzati

CONCORSO N° 12				CONCORSO N° 15			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
40	40	20	X	40	35	25	X
65	20	15	1	35	30	35	X
33	33	33	1	25	35	40	1
50	30	20	1	40	35	25	1
50	30	20	1	35	40	25	X
40	40	20	X	25	50	25	1
33	33	33	1	55	25	20	X
30	40	30	1	55	30	15	1
35	35	30	X	20	40	40	1
60	30	10	X	30	40	30	2
45	30	25	1	25	50	25	1
33	33	33	X	50	30	20	1
50	25	25	1	40	20	40	X
Giocate: 151.789.840 Vincitori: 7.990 Probab. col. vincente 0.0000066222906 N° vinc. "Equi": 1.005				Giocate 160.745.775 Vincitori: 188 Probab. col. vincente 0.0000002165625 N° vinc. "Equi": 34			

Tabella 1. Dati utilizzati

CONCORSO N° 16				CONCORSO N° 17			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
30	50	20	X	30	50	20	2
30	60	10	X	40	50	10	1
40	20	40	X	40	15	45	1
30	25	45	2	40	50	10	1
65	20	15	1	30	20	50	X
30	60	10	2	75	20	5	1
45	10	45	1	50	30	20	X
50	35	15	X	30	60	10	X
35	45	20	1	45	40	15	1
50	40	10	X	50	35	15	1
15	60	25	1	20	60	20	1
20	40	40	X	30	45	25	X
33	33	33	2	15	40	45	2
Giocate: 162.035.960 Vincitori: 61 Probab. col. vincente 0.0000007662154 N° vinc. "Equi": 124				Giocate: 160.417.396 Vincitori: 1.774 Probab. col. vincente 0.0000031492800 N° vinc. "Equi": 505			

Tabella 1. Dati utilizzati

CONCORSO N° 18				CONCORSO N° 20			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
50	30	20	X	15	50	35	X
40	30	30	X	60	35	5	1
30	35	35	X	60	30	10	1
40	35	25	1	33	33	33	2
30	30	40	X	35	50	15	X
25	50	25	2	20	60	20	X
45	35	20	1	45	35	20	2
50	30	20	X	33	33	33	1
25	60	15	X	60	25	15	X
25	45	30	X	45	45	10	1
30	50	20	1	20	55	25	2
50	35	15	1	30	25	45	X
30	45	25	2	15	50	35	X
Giocate: 165.423.481 Vincitori: 64 Probab. col. vincente 0.0000012916968 N° vinc. "Equi": 213				Giocate 163.733.967 Vincitori: 131 Probab. col. vincente 0.0000041347968 N° vinc. "Equi": 677			

Tabella 1. Dati utilizzati

CONCORSO N° 21				CONCORSO N° 22			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
10	40	50	X	30	50	20	1
20	40	40	X	15	50	35	2
50	30	20	1	40	25	35	1
30	45	25	X	30	60	10	X
20	50	30	X	50	45	5	X
40	15	45	1	65	30	5	1
70	25	5	1	50	30	20	X
40	50	10	1	40	30	30	X
25	40	35	2	50	35	15	1
50	35	15	1	50	45	5	1
40	50	10	X	35	50	15	X
30	60	10	X	45	40	15	X
15	35	50	2	45	15	40	1
Giocate: 173.033.517 Vincitori: 3.226 Probab. col. vincente 0.0000529200000 N° vinc. "Equi": 9.156				Giocate: 172.686.230 Vincitori: 1.096 Probab. col. vincente 0.0000149262570 N° vinc. "Equi": 2.577			

Tabella 1. Dati utilizzati

CONCORSO N° 23				CONCORSO N° 25			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
30	25	45	2	35	30	35	1
45	35	20	1	25	25	50	X
35	30	35	X	40	30	30	1
20	40	40	X	40	40	20	X
25	50	25	2	45	30	25	X
55	40	5	1	45	35	20	1
40	35	25	1	35	40	25	1
45	35	20	1	30	50	20	1
50	25	25	X	25	40	35	2
40	35	25	2	45	15	40	X
35	30	35	1	60	25	15	1
33	33	33	X	45	35	20	1
40	25	35	X	15	55	30	1
Giocate: 176.665.259 Vincitori: 570 Probab. col. vincente 0.0000010853841 N° vinc. "Equi": 191				Giocate 170.803.697 Vincitori: 405 Probab. col. vincente 0.0000004219543 N° vinc. "Equi": 72			

Tabella 1. Dati utilizzati

CONCORSO N° 26				CONCORSO N° 27			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
25	35	40	1	40	40	20	X
30	20	50	2	50	30	20	1
20	50	30	1	35	40	25	1
40	15	45	2	40	30	30	1
45	30	25	2	20	60	20	X
55	35	10	1	50	30	20	X
60	25	15	1	40	40	20	X
33	33	33	1	70	20	10	1
30	50	20	X	50	30	20	X
60	25	15	2	40	40	20	1
40	15	45	1	60	30	10	X
30	55	15	2	50	30	20	X
40	35	25	1	40	30	30	X
Giocate: 168.714.875 Vincitori: 38 Probab. col. vincente 0.0000005513062 N° vinc. "Equi": 93				Giocate: 174.580.384 Vincitori: 1.129 Probab. col. vincente 0.0000045722880 N° vinc. "Equi": 798			

Tabella 1. Dati utilizzati

CONCORSO N° 29				CONCORSO N° 30			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
15	50	35	X	35	50	15	1
40	30	20	1	5	35	60	2
60	30	10	1	33	33	33	X
30	50	20	X	33	33	33	X
20	50	30	2	30	50	20	X
20	40	40	X	50	30	20	1
70	20	10	1	50	40	10	X
40	35	25	X	20	50	30	1
50	30	20	X	30	50	20	1
65	30	5	1	50	40	10	X
25	60	15	X	75	20	5	1
10	50	40	X	45	35	20	X
35	40	25	1	20	35	45	2
Giocate: 160.269.050 Vincitori: 9.791 Probab. col. vincente 0.0000361179000 N° vinc. "Equi»: 5788				Giocate 166.942.336 Vincitori: 898 Probab. col. vincente 0.0000064833615 N° vinc. "Equi": 1.082			

Tabella 1. Dati utilizzati

CONCORSO N° 31				CONCORSO N° 32			
Pronostico			Colonna	Pronostico			Colonna
1	X	2	Vincente	1	X	2	Vincente
30	40	15	X	25	50	25	1
50	35	15	X	10	30	60	X
70	25	5	X	30	45	25	1
85	10	5	1	50	30	20	1
25	40	35	1	40	50	10	2
40	40	20	1	40	30	30	X
50	30	20	2	50	30	20	1
60	25	15	1	30	60	10	2
70	20	10	X	33	33	33	X
60	30	10	1	25	60	15	X
30	50	20	2	60	30	10	1
75	20	5	1	33	33	33	1
60	20	20	X	25	40	35	2
Giocate: 165.713.203 Vincitori: 123 Probab. col. vincente 0.0000012852000 N° vinc. "Equi": 212				Giocate: 160.509.948 Vincitori: 748 Probab. col. vincente 0.0000002315486 N° vinc. "Equi": 37			

Tabella 1. Dati utilizzati

CONCORSO N° 34			
Pronostico			Colonna
1	X	2	Vincente
40	50	10	1
60	25	15	1
50	30	20	1
45	35	20	1
80	15	5	1
33	33	33	X
35	50	15	2
45	40	15	1
25	60	15	1
25	40	35	2
50	30	20	X
60	15	25	X
33	33	33	1

Giocate: 150.176.319
Vincitori: 667
Probab. col. vinc.
0.0000012503625
Vincitori "Equi" 187

NOTE:

- Le colonne **PRONOSTICO** sono riprese direttamente dalla "Repubblica".
- La colonna vincente ed il numero di vincitori effettivi sono ripresi direttamente dai dati di **TOTOGUIDA**.
- Il numero di giocate è ottenuto dal montepremi dato da **TOTOGUIDA** diviso per un valore per colonne di 105.64 lire. (Il numero di colonne giocate è pubblicato solo occasionalmente. L'approssimazione è più che accettabile).
- La probabilità della colonna vincente è ottenuta applicando ai risultati della colonna vincente le probabilità del pronostico.
- Il numero di vincitori "Equi" è ottenuto moltiplicando la probabilità della colonna vincente per il numero di giocate.

LISTATI

```
**************************LISTATO DI: IPNUL2***

5 REM "IPNUL"
8 GOSUB 2000
10 ON X GOTO 11,100,1000
11 PRINT "(1) QUESTO PROGRAMMA CONSENTE DI":P
RINT"VERIFICARE SE"
12 PRINT"UN CAMPIONE, DI CUI UNA":PRINT"FRAZI
ONE P1 DI INDIVIDUI";
13 PRINT" POSSIEDE":PRINT"UNA DETERMINATA CAR
ATTERISTICA <Q>,"
14 PRINT"SI POSSA CONSIDERARE APPARTENENTE AD
 UNAPOPOLAZIONE";
15 PRINT" DELLA QUALE SI SUPPONE NOTA":PRINT"
LA PROPORZIONE";
16 PRINT" P2 IN CUI E' PRESENTE LA":PRINT"STE
SSA CARATTERISTICA <Q>."
18 GOSUB 2000
19 ON X GOTO 20,100,1000
20 PRINT"(2) EVIDENTEMENTE E' IMPORTANTE LA N
UME-ROSITA'"
21 PRINT"DEL CAMPIONE , CIOE' IL NUMERO <N> D
I"
22 PRINT"INDIVIDUI CHE LO COMPONGONO."
23 PRINT" SE <N> E' PICCOLO, UNA DIFFERENZA A
NCHEFORTE TRA P1"
24 PRINT"E P2 NON SARA' SUFFICENTE A DIMOSTRA
RE  UNA DIFFERENZA"
25 PRINT"FRA POPOLAZIONE E CAMPIONE; VICEVERS
A SE<N> E' GRANDE"
26 PRINT"UNA DIFFERENZA ANCHE PICCOLA FRA P1
E P2CI PERMETTERA' DI"
27 PRINT"CONCLUDERE CHE IL CAMPIONE NON PUO'"
28 PRINT"PROVENIRE DALLA POPOLAZIONE DATA."
29 GOSUB 2000:ON X GOTO 30,100,1000
30 PRINT"(3) L'ESPRESSIONE SP=SQR(P2*(1-P2)/N
)   CI FORNISCE"
31 PRINT"LA MISURA DELLA DISPERSIONE (LA DEVI
A-  ZIONE STANDARD) CHE"
```

```
32 PRINT"CI POSSIAMO ASPETTARE IN CAMPIONI DI
    GRANDEZZA N"
33 PRINT"ESTRATTI A CASO DA UNA POPOLAZIONE N
ELLAQUALE LA"
34 PRINT"CARATTERISTICA <Q> E' PRESENTE NELLA
    PROPORZIONE P2."
35 PRINT" L'ESPRESSIONE Z1=(P1-P2)/SP CI FORN
ISCEINVECE UNA MISURA"
36 PRINT"STANDARDIZZATA DELLA DIFFERENZA CHE
    ABBIAMO RISCONTRATO"
37 PRINT"FRA IL NOSTRO CAMPIONE E LA POPOLAZI
ONE NOTA."
38 GOSUB 2000
39 ON X GOTO 40,100,1000
40 PRINT"(4) LA GRANDEZZA <Z1> VIENE CONFRONT
ATA CON UNA TABELLA"
41 PRINT"STANDARD, MEMORIZZATA MEDIANTE 'DATA
'    NEL PROGRAMMA,"
42 PRINT"CHE RAPPRESENTA LA DISTRIBUZIONE NOR
MALEDI PROBABILITA' (GAUSS)."
43 PRINT" DA QUESTA RICAVIAMO LA PROBABILITA'
 CHE";
44 PRINT "UNO SCOSTAMENTO UGUALE A QUELLO TRO
VATO POSSA VERIFICARSI PER CASO."
45 PRINT" L'IPOTESI NULLA E' CHE CAMPIONE E P
OPOLAZIONE SIANO OMOGENEI, E"
46 PRINT"SI CONSIDERA <RESPINTA> SE LA PROBAB
I-   LITA'DI OTTENERE PER CASO"
47 PRINT"UNO SCOSTAMENTO ALMENO PARI A QUELLO
    OTTENUTO E'"
48 PRINT"INFERIORE A 0.005 (0.5 %)."
50 GOSUB 2000
55 ON X GOTO 11,100,1000
100 DIM Z(35),A(35)
110 PRINT" "
120 PRINT"INIZIO"
130 PRINT" "
135 PRINT"ATTENZIONE, NON USARE PERCENTUALI :
     50% VA INDICATO COME 0.5"
```

```
140 PRINT" "
199 REM LEGGE TABELLA DISTRIBUZIONE NORMALE
200 DATA 0,0,0.1,0.0396,0.2,0.0791,0.3,0.1180
,0.4,0.1556
202 DATA 0.5,0.1917,0.6,0.2259,0.7,0.2581,0.8
,0.2881
204 DATA 0.9,0.3158,1.0,0.3411,1.1,0.3641,1.2
,0.38847,1.3,0.4031
206 DATA 1.4,0.4192,1.5,0.4333,1.6,0.4454,1.7
,0.4556,1.8,0.4643
208 DATA 1.9,0.4715,2.0,0.4774,2.1,0.4823,2.2
,0.4862,2.3,0.4893
209 DATA 2.4,0.4917,2.5,0.4937,2.6,0.4952,2.7
,0.4963,2.8,0.4972,2.9,0.4979
210 DATA 3.0,0.4984,3.1,0.4988,3.2,0.4991,3.3
,0.4993,3.4,0.4995,3.5,0.4996
240 RESTORE
250 FOR I=0 TO 35
260 READ Z(I),A(I)
270 NEXT I
399 REM INPUT DATI
400 PRINT "PROPORZIONE DELLA CARATTERISTICA N
ELLA  POPOLAZIONE P2=":INPUT P2
410 PRINT "PROPORZIONE DELLA CARATTERISTICA N
EL     CAMPIONE P1=":INPUT P1
420 PRINT "NUMERO DI INDIVIDUI NEL CAMPIONE N
=":INPUT N
499 REM CALCOLO
500 SP=SQR(P2*(1-P2)/N)
510 Z1=(P1-P2)/SP
520 Z1=ABS(Z1)
599 REM INTERPOLA
600 FOR I=0 TO 35
610 IF Z1>Z(I) GOTO 650
620 R1=A(I)+(A(I1)-A(I))*(Z1-Z(I))/(Z(I1+1)-Z
(I))
640 GOTO 680
650 NEXT I
660 R1=A(I-1)
680 R1=.5-R1
```

```
630 PRINT"UN CAMPIONE COME QUELLO OTTENUTO  P
OTREBBE ESSERE OTTENUTO"
691 PRINT"PER CASO DALLA POPOLAZIONE INDICATA
 CON ";
700 PRINT "PROBABILITA'=";R1
800 IF R1>=0.005 GOTO 900
810 PRINT"L'IPOTESI NULLA E' RESPINTA : UN CA
MPIONE COME QUELLO"
820 PRINT"INDICATO NON APPARTIENE ALLA POPOLA
ZIONE INDICATA"
830 GOTO 1000
900 PRINT"L'IPOTESI NULLA NON PUO' ESSERE RES
PINTA"
910 PRINT"IL CAMPIONE PUO' APPARTENERE AD UNA
 POPOLAZIONE UGUALE"
920 PRINT"ALLA POPOLAZIONE INDICATA"
1000 STOP
2000 PRINT" "
2005 INPUT"(1)=HELP (2)=ESEGUI (3)=STOP:";X
2007 PRINT" "
2010 ON X GOTO 2030,2030,2030
2020 GOTO 2000
2030 RETURN
```

************************LISTATO DI: CONTGZ*****

```
10 REM "CONTGZ"
20 X=0
22 GOSUB 2000
24 ON X GOTO 30,100,1000
30 PRINT"(1) QUESTO PROGRAMMA VERIFICA LA IND
I-"
31 PRINT"PENDENZA FRA DUE FATTORI DI CLASSI-"
32 PRINT"FICAZIONE DI UN GRUPPO DI INDIVIDUI.
"
33 PRINT"LA DOMANDA A CUI DOBBIAMO RISPONDE-"
34 PRINT"RE E' SE QUESTI DUE FATTORI POSSANO"
35 PRINT"ESSERE CONSIDERATI INDIPENDENTI."
36 PRINT"GLI INDIVIDUI POSSONO AD ESEMPIO ESS
ERE PERSONE ED"
```

```
37 PRINT"I FATTORI DI CLASSIFICAZIONE ESSERE
LA  STATURA ED IL PESO"
40 GOSUB 2000
42 ON X GOTO 44,100,1000
44 PRINT"(2) COSTRUIAMO UNA TABELLA NELLA QUA
LE LE"
45 PRINT"RIGHE RAPPRESENTANO IL PRIMO FATTORE
"
46 PRINT"DI CLASSIFICAZIONE E LE COLONNE RAPP
RE-"
47 PRINT"SENTANO IL SECONDO FATTORE DI CLASSI
FI-"
48 PRINT"CAZIONE."
49 PRINT"OGNI RIGA RAPPRESENTA UN DIVERSO LIV
EL-"
50 PRINT"LO DEL PRIMO FATTORE E RISPETIVAMENT
E"
51 PRINT"OGNI COLONNA RAPPRESENTA UN DIVERSO"
52 PRINT"LIVELLO DEL SECONDO."
53 PRINT"NELLA GENERICA CASELLA ALLA I-ESIMA"
54 PRINT"RIGA E J-ESIMA COLONNA VIENE INSERIT
O"
55 PRINT"IL NUMERO N(I,J) DI CASI CHE SONO ST
ATI"
56 PRINT"CLASSIFICATI AL LIVELLO I-ESIMO RIGU
ARDO"
57 PRINT"AL PRIMO FATTORE ED AL LIVELLO J-ESI
MO"
58 PRINT"RIGUARDO AL SECONDO FATTORE"
60 GOSUB 2000
62 ON X GOTO 70,100,1000
70 PRINT"(3) IL PROGRAMMA CALCOLA LA FUNZIONE
"
72 PRINT"STATISTICA CHI-QUADRO PER LA TABELLA
"
74 PRINT"CARICATA E CONFRONTA IL VALORE TROVA
TO"
76 PRINT"CON I VALORI TABULATI DI CHI-QUADRO"
78 PRINT"AL LIVELLO DI PROBABILITA' 0.05."
```

```
80 PRINT"LA TABELLA CHI-QUADRO E' CONTENUTA"
82 PRINT" IN UN DATA NEL PROGRAMMA."
84 PRINT"LA TABELLA MEMORIZZATA SI ARRESTA"
86 PRINT"AD UN VALORE DI (NR-1)*(NC-1)=10"
87 PRINT"IN CUI NR=NUMERO DI RIGHE DELLA TABE
LLA"
88 PRINT"ED NC=NUMERO DI COLONNE."
90 GOSUB 2000
92 ON X GOTO 93,100,1000
93 PRINT"(4) L'IPOTESI NULLA E' CHE I DUE FAT
TORI SIANO <INDIPENDENTI>"
94 PRINT"RESPINGERE L'IPOTESI NULLA PORTA A C
ONCLUDERE CHE I DUE FATTORI"
95 PRINT"NON SONO INDIPENDENTI , CIOE' CHE UN
O DETERMINA L'ALTRO"
96 PRINT"OPPURE CHE ENTRAMBI SONO DETERMINATI
 DA UN TERZO FATTORE COMUNE"
97 GOSUB 2000
98 ON X GOTO 30,100,1000
100 REM ESEGUI
105 GOSUB 3000
110 INPUT"FATTORE A) NUMERO DI RIGHE NR=";NR
120 INPUT"FATTORE B) NUMERO COLONNE  NC=";NC
130 DIM N(NR,NC)
135 DIM P(NR,NC)
140 DIM O(NR,NC)
150 FOR I=0 TO NR
155 FOR J=0 TO NC
160 N(I,J)=0
165 NEXT J
170 NEXT I
200 FOR I=1 TO NR
210 FOR J=1 TO NC
220 PRINT"FATTORE A) LIVELLO:";I
230 PRINT"FATTORE B) LIVELLO:";J
240 INPUT"NUMERO DI CASI N(I,J)=";N(I,J)
250 NEXT J
260 NEXT I
300 REM SOMME
```

```
310 FOR I=1 TO NR
320 FOR J=1 TO NC
330 N(I,0)=N(I,0)+N(I,J)
340 N(0,J)=N(0,J)+N(I,J)
350 N(0,0)=N(0,0)+N(I,J)
360 NEXT J
370 NEXT I
400 FOR I=1 TO NR
410 P(I,0)=N(I,0)/N(0,0)
420 NEXT I
430 FOR J=1 TO NC
440 P(0,J)=N(0,J)/N(0,0)
450 NEXT J
460 FOR I=1 TO NR
470 FOR J=1 TO NC
480 O(I,J)=N(I,J)/N(0,0)
490 P(I,J)=P(I,0)*P(0,J)
500 NEXT J
510 NEXT I
580 CHI2=0
600 FOR I=1 TO NR
610 FOR J=1 TO NC
620 CHI2=CHI2+((P(I,J)-O(I,J))↑2/P(I,J))
630 NEXT J
640 NEXT I
660 CHI2=CHI2*N(0,0)
700 IF CHI2>CHI2((NR-1)*(NC-1)) GOTO 800
720 PRINT"NON ESISTE EVIDENZA DI DIPENDENZA T
RA I FATTORI"
730 GOTO 1000
800 PRINT"ESISTE EVIDENZA DI UNA DIPENDENZA T
RA I FATTORI"
1000 STOP
2000 PRINT" "
2005 INPUT "(1)=HELP (2)=ESEGUI (3)=STOP:";X
2007 PRINT" "
2010 ON X GOTO 2030,2030,2030
2020 GOTO 2000
2030 RETURN
3000 REM TAVOLA CHI-QUADRO
```

```
3005 DIM CHI2(10)
3010 DATA 3.841,5.991,7.815,9.488,11.070,12.5
92
3020 DATA 14.067,15.507,16.919,18.307
3030 RESTORE
3040 FOR I=1 TO 10
3050 READ CHI2(I)
3060 NEXT I
3090 RETURN
```

```
*********************LISTATO DI: TOTOBRIER2*******
```

```
10 REM TOTOBRIER
20 GOSUB 2000
25 ON X GOTO 30,100,1000
30 REM
32 PRINT"(1) QUESTO PROGRAMMA CALCOLA IL":PRI
NT"VOTO-ERRORE"
34 PRINT"MEDIO SU UNA SERIE DI CONCORSI."
36 PRINT" INSERIRE LA COLONNA VINCENTE E"
38 PRINT"SUCCESSIVAMENTE IL PRONOSTICO PER LE
 13"
40 PRINT"PARTITE."
42 PRINT" LA PROBABILITA' SOMMA AD 1 , PERTAN
TO"
44 PRINT"UNA PROBABILITA' AD ESEMPIO DEL 50%"
46 PRINT"VA INDICATA CON 0.5; IL 33% VA INDIC
ATO"
48 PRINT"CON .3333333. MAGGIORE E' IL NUMERO"
50 PRINT"DI DECIMALI E PIU' PRECISO E' IL":PR
INT"RISULTATO."
60 GOSUB 2000
62 ON X GOTO 30,100,1000
100 DIM X$(13),Y$(1)
102 DIM P(3)
105 TG=0.577350 : CS=1.154734
110 REM  IN TG E CS I VALORI DELLA TANGENTE E
 IL RECIPROCO DEL COSENO DI 30 G.
```

```
120 ET=0 : NP=0
130 EM=0
250 INPUT"COLONNA VINCENTE F=FINE:";X$
270 IF X$="F" GOTO 3000
300 FOR I=1 TO 13
310 PRINT"PARTITA:";I
320 INPUT"PRONOSTICO  P1,PX,P2:";P(1),P(2),P(
3)
330 GOSUB 4000
340 NEXT I
350 GOTO 250
1000 STOP
2000 PRINT" "
2005 INPUT"(1)=HELP (2)=ESEGUI (3)=STOP:";X
2007 PRINT" "
2010 ON X GOTO 2030,2030,2030
2020 GOTO 2000
2030 RETURN
3000 REM RISULTATI
3010 PRINT"ERRORE MEDIO:";EM
3015 PRINT" "
3020 IF EM>.43 THEN IF EM<.45 THEN PRINT"PRON
OSTICATORE <NULLO>"
3030 IF EM>.45 THEN  PRINT"QUI' SI TIRA A IND
OVINARE"
3040 IF EM<=.43 THEN IF EM>.415 THEN PRINT"CO
MINCIAMO A VEDERE QUALCHE COSA"
3050 IF EM<=.415 THEN IF EM>.4 THEN  PRINT"NO
N C'E' MALE: INSISTI"
3060 IF EM<=.4 THEN IF EM>.375 THEN PRINT"MOL
TO BENE, VERAMENTE INTERESSANTE"
3070 IF EM<=.375 THEN IF EM>.35 THEN  PRINT"E
CCEZIONALE, SE NON E' UN CASO"
3080 IF EM<=.35 THEN IF EM>.33333333 THEN PRI
NT"MI LASCI PERPLESSO..."
3090 IF EM<=.3333333 THEN PRINT"TU BARI!! SE
NO PUOI SPERARE NEL NOBEL"
3200 STOP
4000 REM TOTALIZZA
```

```
4010 PA=P(2) : PB=P(3)
4020 IF MID$(X$,I,1)="X" THEN PA=P(1)
4030 IF MID$(X$,I,1)="2" THEN PB=P(1)
4050 GOSUB 4500
4070 ET=ET+ER : NP=NP+1
4080 EM=ET/NP
4120 RETURN
4500 REM ERR
4510 ER=PA↑2+(PA*TG+PB*CS)↑2
4520 RETURN
```

```
*************************LISTATO DI: TOTOFIN2****
```

```
10 REM TOTOFIN2 PROGRAMMA GENERALE
11 REM PROGRAMMA COMPLETO
20 DIM P1(13,3),CC(13,1)
22 REM IN P1 VANNO LE PROBABILITA' DEI RISULT
ATI DEL CONCORSO DA GIOCARE
23 REM CHE VENGONO PRESE IN INPUT ALLA 100-13
0
24 REM IN CC VA LA CURVA C LETTA DAL DATA ALL
A 9065
25 REM CHE DA' COPPIE DI VALORI COLONNA PROGR
ESSIVA-NUMERO DI PUNTATE
28 DIM CM(40,1),R1(200),TT(2),RF(25)
30 REM IN CM VA LA CURVA CALCOLATA NUMERO DI
SCHEDE-PROBABILITA'
31 REM AD INTERVALLI DI PROBABILITA' CONTROLL
ATI ALLA 1050
32 REM IN R1 VANNO LE COLONNE ACCETTATE COME
NUMERI TERNARI
33 REM IN TT VNNO LE PROBAB.DELLA TROTTOLA,SE
TTATE ALLA 50
34 REM IN RF VANNO I VALORI,LETTI NEL DATA,DE
LLE DIFFERENZE FRA NUMERI TERNARI
35 REM CHE RAPPRESENTANO COLONNE CON UN SOLO
SEGNO DI DIFFERENZA
```

```
40 NR=9 : NC=21
41 DIM TS(NR,NC)
42 REM IN TS VENGONO MEMORIZZATE IN ORDINE PR
OGRESSIVO
43 REM LE COLONNE GENERATE A CASO PER COSTRUI
RE POI LA CURVA CM
44 REM NR E NC CONTROLLANO IL NUMERO DI COLON
NE GENERATE
45 DIM C(12,2),ST(12),FW(12)
46 REM C,ST,FR SERVONO A PRODURRE LE COLONNE
48 DIM Z$(13),X$(1)
49 REM IN GENERALE GLI INDICI CORRISPONDONO A
I RISULTATI 0=1 1=X 2=2
50 TT(0)=.5 : TT(1)=.3
51 TT(2)=.2
60 TS(0,0)=-1 : TS(0,1)=-1
70 PRINT"ATTENZIONE LE PROBABILITA' NON VANNO
"
72 PRINT"ESPRESSE IN PERCENTUALE ."
73 PRINT"AD ESEMPIO 50% VA INDICATO CON 0.5"
100 FOR I=1 TO 13
105 PRINT"PARTITA:";I
110 INPUT"PROBABILITA' P1,PX,P2:";P1(I,0),P1(
I,1),P1(I,2)
130 NEXT I
150 RESTORE
155 REM READ DIFFERENZE PER RIDUZIONE
160 FOR I=0 TO 25
165 READ RF(I)
170 NEXT I
260 REM READ CURVA C
270 FOR I=0 TO 13
274 READ CC(I,0),CC(I,1)
280 NEXT I
300 PRINT"INIZIA ROUTINE MONTECARLO"
310 PRINT"CIRCA ";INT((NR*NC)*3/4);" COLONNE"
400 REM
401 REM
402 REM TEMPI DI ESECUZIONE E VISUALIZZAZIONE
 DI CONTROLLO
```

```
403 REM LA ROUTINE MONTECARLO PRODUCE COLONNE
 E LE INSERISCE NELLA
404 REM MATRICE DI SORT ALLA VELOCITA' DI CIR
CA 80 COLONNE AL MINUTO
405 REM  (OVVIAMENTE SU QUESTA MACCHINA)
406 REM IL NUMERO TOTALE DI COLONNE GENERATE
E' DI CIRCA TRE QUARTI
407 REM DELLA CAPACITA' TOTALE DELLA MATRICE
DIMENSIONATA ALLA 40
408 REM CON UN DIMENSIONAMENTO 49*51=2000 CIR
CA SONO RICHIESTI
409 REM CIRCA 20 MINUTI PER LA PREPARAZIONE D
EI DATI
410 REM IL SISTEMA VISUALIZZA IL PROGRESSIVO
DI COLONNA GENERATA
411 REM
412 REM
470 REM GENERA A CASO UN CAMPIONE DI COLONNE
471 REM NE CALCOLA LA PROBAB. E MEMORIZZA
472 REM LE PROBAB. IN TS ORDINATE PER VALORI
DECRESCENTI
473 REM OGNI RIGA DELLA MATRICE TS CONTIENE
474 REM IN COLONNA 0 IL NUMERO INDICE DELLA R
IGA CHE CONTIENE
475 REM  I VALORI DI PROBAB. IMMEDIATAMENTE P

IU' ALTI
476 REM ED IN COLONNA 1 L'INDICE DELLA RIGA D
EI VALORI IMMEDIATAM PIU' BASSI
477 REM -1 E' IL TAPPO,IN OGNI RIGA DA COLONN
A 2 SONO CONTENUTI I VALORI
480 NT=2000
500 FOR KK=1 TO NT
580 REM PS=PROBAB. DELLA SCHEDA GENERATA
585 REM GENERA A CASO
590 PS=1
600 FOR J=1 TO 13
610 X=INT(RND(1)*3)
620 PS=PS*P1(J,X)
650 NEXT J
```

```
740 GOSUB 7000
745 IF ER>0 GOTO 900
750 PRINTKK
800 NEXT KK
900 NT=KK
920 PRINT"INIZIA COSTRUZIONE CURVA"
930 RP=INT(1594323/NT):SC=1
940 W1=1
945 GOSUB 1200
950 PX=P0
951 REM PX=PROBAB.DELLA COLONNA PIU' PROBABIL
E
960 TC=0 : CM(0,0)=PX
962 CM(0,1)=SC
1000 REM LEGGENDO IN TS COSTRUISCE LA CURVA I
N CM
1010 P0=PX : I=0
1020 FOR J=2 TO NC
1025 PR=TS(I,J)
1030 IF PR=0 GOTO 1100
1032 IF W1<>0 GOTO 1040
1036 W1=1
1040 SC=SC+RP
1045 IF PR<5.0E-7 GOTO 1140
1050 IF P0-PR<5.0E-7 GOTO 1100
1060 TC=TC+1 : CM(TC,0)=PR
1062 CM(TC,1)=SC : P0=PR
1070 W1=0
1100 NEXT J
1110 I=TS(I,0)
1120 IF I=-1 GOTO 1140
1130 GOTO 1020
1140 W1=0
1145 GOSUB 1200
1150 PY=P0 : TC=TC+1
1152 CM(TC,0)=PY : CM(TC,1)=1594323
1155 REM PY=PROBAB. DELLA COLONNA MENO PROBAB
ILE
1160 GOTO 2090
```

```
1200 REM CALCOLA LA PROBABILITA' DELLA COLONN
A PIU' E MENO PROBABILE
1205 PS=1
1220 FOR I=1 TO 13
1230 P0=0
1235 IF W1=0 THEN P0=1
1240 FOR J=0 TO 2
1250 IF W1=1 THEN IF P0<P1(I,J) THEN P0=P1(I,
J)
1255 IF W1=0 THEN IF P0>P1(I,J) THEN P0=P1(I,
J)
1260 NEXT J
1290 PS=PS*P0
1310 NEXT I
1320 P0=PS
1330 RETURN
2090 REM
2100 GOTO 2991
2360 REM ROUTINE DI CONTROLLO ACCETTAZIONE DE
LLA COLONNA
2370 GOSUB 4800
2375 IF PS<AP GOTO 2540
2380 GOSUB 4400
2385 GOSUB 4600
2388 FF=FQ*FT/100
2398 IF PS/FF<AT GOTO 2550
2400 REM CONTROLLA RIDUZIONE
2405 IF RI=0 GOTO 2455
2410 FOR F=1 TO R1
2415 DF=CL-R1(F)
2420 FOR JK=0 TO 25
2425 IF DF=RF(JK) GOTO 2560
2430 NEXT JK
2440 NEXT F
2455 PRINT"OK"
2500 R1=R1+1 : R1(R1)=CL
2520 AX=AX+1
2521 REM GOTO 2530
2522 PRINT"TOTALE RICHIESTE:";T0;"  GENERATE:
";NN
```

```
2525 PRINT"ACCETTATE:";AX
2526 PRINT"SCARTATE PER BASSA PROBABILITA':";
S9
2527 PRINT"SCARTATE PER RIDUZIONE SISTEMA:";S
8
2528 PRINT"SCARTATE PER BASSO RAPPORTO TOTOMA
C:";S7
2530 RETURN
2540 S9=S9+1
2545 RETURN
2550 S7=S7+1
2555 RETURN
2560 S8=S8+1
2565 RETURN
2800 REM
2900 REM ROUTINE EVENTUALE PER GENERAZIONE CO
LONNE A CASO
2905 REM NON UTILIZZATA IN QUESTO PROGRAMMA
2910 FOR I=0 TO 12
2915 ST(I)=INT(RND(1)*3)
2920 NEXT I
2930 GOSUB 2360
2980 GOTO 2910
2990 STOP
2991 REM SVILUPPO COLONNE RICHIESTE
2992 REM LA ROUTINE DI SVILUPPO COLONNARE PRO
DUCE CIRCA 200 COLONNE AL MINUTO
2993 REM IL CONTROLLO DI ACCETTAZIONE,SENZA L
A RIDUZIONE DEL SISTEMA , A CIRCA
2994 REM 55 COLONNE AL MINUTO . LE COLONNE PR
ODOTTE VENGONO VISUALIZZATE
2995 REM COME NUMERO TERNARIO (ISTRUZ.3287).
2996 REM IL TOTALE DI COLONNE PRODOTTE E SCAR
TATE PER LE DIVERSE CAUSE
2997 REM E' VISUALIZZATO (ISTRUZ.2522-2528) ,
 LA VISUALIZZAZIONE RALLENTA
2998 REM IL PROCESSO A CIRCA 42 COLONNE/MINUT
O . PUO' ESSERE ELIMINATA COL
2999 REM GOTO ALLA 2521
```

```
3000 REM LA RIDUZIONE DEL SISTEMA RALLENTA SE
NSIBILMENTE IL PROCESSO
3001 REM IL NUMERO EFFETTIVO DI COLONNE ACCET
TATE DIPENDE DALLA GRIGLIA
3002 PRINT"INSERITE I SEGNI DESIDERATI PER"
3003 PRINT"CIASCUNA PARTITA"
3005 REM QUESTO PROGRAMMA NON UTILIZZA PRONOS
TICI SEPARATI
3006 REM COME PRONOSTICO <PERSONALE> VIENE US
ATO LO STESSO
3007 REM PRONOSTICO USATO PER COSTRUIRE LA CU
RVA CM
3010 FOR I=0 TO 12
3012 PRINT"PARTITA:";I+1
3014 INPUT"1  0=NO 1=SI:";C(I,0)
3016 INPUT"X  0=NO 1=SI:";C(I,1)
3018 INPUT"2  0=NO 1=SI:";C(I,2)
3020 NEXT I
3022 PRINT"GRIGLIA DI SELEZIONE"
3024 INPUT"RIDUZIONE 0=NO 1=SI:";RI
3026 INPUT"PROBAB.MINIMA ACCETTATA:";AP
3028 INPUT"RAPPORTO TOTOMAC MINIMO:";AT
3030 AX=0 : S7=0
3032 S8=0 : S9=0
3040 REM CONTEGGIO
3045 T0=1
3050 I=0
3055 T1=0
3060 FOR J=0 TO 2
3065 T1=T1+C(I,J)
3070 NEXT J
3075 T0=T0*T1
3080 IF I=12 GOTO 3085
3082 I=I+1
3083 GOTO 3055
3085 PRINT"TOTALE DA GENERARE:";T0
3090 INPUT "CR:";BL
3095 REM PRIMA COLONNA
3100 CL=0
```

```
3105 FOR I=0 TO 12
3110 FW(I)=-1
3115 FOR J=0 TO 2
3120 IF C(I,J)=0 GOTO 3140
3125 CL=CL+3↑I*J
3130 IF FW(I)>=0 GOTO 3135
3132 FW(I)=J
3133 ST(I)=J
3135 GOTO 3145
3140 NEXT J
3145 NEXT I
3150 NN=1
3165 LP=CL
3167 GOSUB 2360
3170 REM GENERA
3175 IK=0
3180 GOSUB 3215
3185 IF W1>0 THEN IF IK=12 THEN GOTO 5000
3190 IF W1>0 GOTO 3205
3195 IF IK=0 GOTO 3180
3200 GOTO 3175
3205 IK=IK+1
3210 GOTO 3180
3215 REM
3220 W1=1
3225 IF IK=0 GOTO 3235
3230 GOSUB 3335
3235 GOSUB 3265
3240 IF CL=LP GOTO 3255
3245 LP=CL
3250 W1=0
3255 RETURN
3265 REM
3270 CL=LP
3275 GOSUB 3300
3280 IF CL=LP THEN RETURN
3285 NN=NN+1
3287 PRINT"TERNARIO COLONNA";CL
3290 GOSUB 2360
```

```
3295 RETURN
3300 REM AVANZA COLONNA
3302 GOSUB 3700
3303 IF ST(IK)=2 GOTO 3330
3305 FOR J=ST(IK)+1 TO 2
3310 IF C(IK,J)=0 GOTO 3325
3315 CL=CL+KW*(J-ST(IK)) : ST(IK)=J
3320 IF CL<>LP GOTO 3330
3325 NEXT J
3330 RETURN
3335 REM RESET COLONNA
3337 GOSUB 3700
3340 CL=INT((CL)/KW)   : CL=CL*KW
3345 FOR J=0 TO IK-1
3350 ST(J)=FW(J)
3355 CL=CL+3↑J*FW(J)
3360 NEXT J
3365 LP=CL
3370 RETURN
3700 REM
3710 KW=1
3720 IF IK=0 THEN RETURN
3730 FOR KV=1 TO IK
3740 KW=KW*3
3750 NEXT KV
3760 RETURN
4000 STOP
4400 REM TROVA LA GAMMA SULLA CURVA C
4410 FOR I=0 TO TC
4420 IF PS>CM(I,0) GOTO 4460
4430 QP=CM(I,0) : QS=CM(I,1)
4440 NEXT I
4460 ZP=CM(I,0) : ZS=CM(I,1)
4470 X0=(QP-PS)/(QP-ZP)
4490 UM=QS+(ZS-QS)*X0
4520 RETURN
4600 REM FREQUENZA CURVA C
4610 FOR I=0 TO 13
4620 IF CC(I,0)>UM GOTO 4670
```

```
4630 QP=CC(I,0) : QS=CC(I,1)
4640 NEXT I
4650 STOP
4670 ZP=CC(I,0) : ZS=CC(I,1)
4680 X0=(UM-QP)/(ZP-QP)
4700 FQ=QS+(ZS-QS)*X0
4750 RETURN
4800 REM CALCOLO PROBABILITA' E TROTTOLA
4810 PS=1 : FT=1
4820 FOR JK=1 TO 13
4850 PS=PS*P1(JK,ST(JK-1))
4870 FT=FT*TT(ST(JK-1))
4880 NEXT JK
4900 RETURN
5000 REM DECODIFICA E STAMPA
5010 PRINT"QUESTO PROGRAMMA STAMPA LE COLONNE
"
5011 PRINT"ACCETTATE A VIDEO ."
5012 PRINT"PER AVERE UN ALTRO TIPO DI STAMPA"
5013 PRINT"MODIFICARE IL LISTING ALLA 5450"
5014 INPUT"BATTERE RETURN PER CONTINUARE";BL
5030 FOR J=1 TO R1
5040 X=R1(J): Z$="             "
5100 FOR I=12 TO 0 STEP -1
5110 M1%=(X+.1)/3↑I
5120 ON M1%+1 GOTO 5200,5240,5280
5130 STOP
5200 X$="1"
5210 GOTO 5300
5240 X$="X"
5250 GOTO 5300
5280 X$="2"
5300 Z$=LEFT$(Z$,I)+X$+MID$(Z$,I+2)
5320 X=X-M1%*3↑I
5400 NEXT I
5450 PRINT"COLONNA:";Z$
5455 INPUT "CR";BL
5500 NEXT J
5600 PRINT"FINE PROGRAMMA"
```

```
5700 STOP
7000 REM SORT IN MATRICE
7010 I=0 : ER=0
7020 IF TS(I,0)=-1 GOTO 7300
7030 IF TS(I,2)=PS GOTO 7300
7040 IF TS(I,2)<PS GOTO 7100
7042 IF TS(I,0)=-1 GOTO 7300
7044 IF TS(TS(I,0),2)<=PS GOTO 7300
7050 I=TS(I,0)
7060 GOTO 7030
7100 I=TS(I,1)
7105 IF I<0 THEN I=0
7110 GOTO 7300
7300 REM INSERT
7330 IF TS(I,NC)<>0 GOTO 7600
7350 FOR J=NC TO 2 STEP -1
7360 IF TS(I,J)=0 GOTO 7490
7370 IF PS<=TS(I,J) GOTO 7500
7380 TS(I,J+1)=TS(I,J)
7490 NEXT J
7492 J=2
7495 J=J-1
7500 TS(I,J+1)=PS
7550 RETURN
7600 REM SPEZZA LA RIGA
7610 FOR K=00 TO NR
7620 IF TS(K,0)=0 GOTO 7650
7630 NEXT K
7635 ER=1
7640 RETURN
7650 REM
7652 TS(K,0)=TS(I,0) : TS(I,0)=K
7653 TS(K,1)=I
7655 IF TS(K,0)>=0 THEN  TS(TS(K,0),1)=K
7680 FOR J=2 TO NC/2+1
7690 TS(K,J)=TS(I,J+NC/2) : TS(I,J+NC/2)=0
7700 NEXT J
7710 GOTO 7010
7999 STOP
```

```
8000 INPUT"N:";N
8005 X=3
8010 FOR I=2 TO N
8020 X=X*3
8030 NEXT I
8040 PRINTZ$
8050 PRINTX
8060 GOTO 8000
9050 DATA 1,2,3,6,9,18,27,54,81,162,243,486,7
29
9055 DATA 1458,2187,4374,6561,13122,19683,393
36,59049,118098
9060 DATA 177147,354294,531441,1062882
9065 DATA 1,137330,3,72000,10,34670,30,18000
9070 DATA 100,8670,300,4400,1000,2130,3000,11
00
9075 DATA 10000,520,30000,273,100000,130
9080 DATA 300000,67,1000000,33,1600000,26,0,0
44500 STOP
```

```
*******************LISTATO DI: TOTOTARA2********

10 REM "TOTOTARA"
20 GOSUB 2000
25 ON X GOTO 30,100,1000
30 PRINT"(1) QUESTO PROGRAMMA RICEVE IN INPUT
 LE"
32 PRINT"COLONNE VINCENTI DI UN NUMERO QUALSI
ASI"
34 PRINT"DI CONCORSI ED I PRONOSTICI RELATIVI
"
36 PRINT"ESPRESSI DALLA PROBABILITA' DEI RISU
L-"
38 PRINT"TATI 1,X,2."
40 PRINT"AL TERMINE DEL CARICAMENTO, IL PROGR
AM-"
42 PRINT"MA CALCOLA LA FREQUENZA CON CUI SI S
ONO"
```

```
44 PRINT"VERIFICATI I RISULTATI E LA CONFRONT
A"
46 PRINT"CON LA PROBABILITA' ASSEGNATA AD ESS
I"
48 PRINT"DAL PRONOSTICATORE , FORNENDONE QUIN
DI"
50 PRINT"UNA TARATURA ."
60 GOSUB 2000
62 ON X GOTO 70,100,1000
70 PRINT"(2) SONO AMMESSI NEL PRONOSTICO VALO
RI  DI"
72 PRINT"PROBABILITA' 5-10-15-20 ETC. FINO A
100"
74 PRINT"ED IL VALORE 33."
76 PRINT" LA COLONNA VINCENTE VA CARICATA NEL
LA"
78 PRINT"FORMA STANDARD ES.:11X1X2111X21X"
80 GOSUB 2000
82 ON X GOTO 30,100,1000
100 DIM X$(13),Y$(1)
102 DIM F(30,1),P(3)
110 DATA 1,5,2,10,3,15,4,20,5,25,6,30,7,33,8,
35,9,40
120 DATA 10,45,11,50,12,55,13,60,14,65,15,70,
16,75
130 DATA 17,80,18,85,19,90,20,95,21,100,0,0
150 FOR I=1 TO 30
160 FOR J=0 TO 1
170 F(I,J)=0
180 NEXT J
190 NEXT I
250 INPUT"COLONNA VINCENTE F=FINE:";X$
270 IF X$="F" GOTO 3000
300 FOR I=1 TO 13
310 PRINT"PARTITA:";I
320 INPUT"PRONOSTICO  P1,PX,P2:";P(1),P(2),P(
3)
330 GOSUB 4000
335 IF N=0 GOTO 310
```

```
340 NEXT I
350 GOTO 250
1000 STOP
2000 PRINT" "
2005 INPUT"(1)=HELP (2)=ESEGUI (3)=STOP:";X
2007 PRINT" "
2010 ON X GOTO 2030,2030,2030
2020 GOTO 2000
2030 RETURN
3000 REM RISULTATI
3010 FOR I=1 TO 21
3020 RESTORE
3030 READ M,N
3040 IF M=0 THEN STOP
3050 IF M<>I GOTO 3030
3060 IF F(I,0)=0 GOTO 3500
3070 Y=F(I,1)*100/F(I,0)
3080 PRINT"PRONOSTICO:";N
3090 PRINT"ASSEGNATO:";F(I,0);"  USCITO:";F(I,1)
3100 PRINT"PERCENTUALE EFFETTIVA:";Y
3110 INPUT"RETURN PER CONTINUARE:";BL
3120 NEXT I
3200 STOP
3500 Y=-1
3510 GOTO 3080
4000 REM TOTALIZZA
4010 X=1
4020 IF MID$(X$,I,1)="X" THEN X=2
4030 IF MID$(X$,I,1)="2" THEN X=3
4040 FOR J=1 TO 3
4050 RESTORE
4060 READ M,N
4070 IF N=0 THEN GOTO 4120
4080 IF N<>P(J) GOTO 4060
4090 F(M,0)=F(M,0)+1
4100 IF X=J THEN F(M,1)=F(M,1)+1
4110 NEXT J
4120 RETURN
```

L'USO PRATICO DEL PROGRAMMA

Finora abbiamo trattato il problema del calcolo della schedina ottimale, privilegiando gli aspetti teorici, in modo da permettere a chiunque di avere piena coscienza dei principi di funzionamento di questo sistema software. Tuttavia questo non è il solo obbiettivo che si intende perseguire, ma vogliamo anche fornire un buon metodo per il calcolo delle più convenienti colonne da giocare. Per raggiungere questo scopo ci è sembrato utile riprendere l'intera materia trattata, approfondendo gli aspetti relativi all'interazione con il calcolatore.

Nel seguito verrà brevemente richiamato il problema che ogni modulo software risolve e, con maggiore dettaglio, le modalità di utilizzo del modulo stesso.

Utilizzo del programma IPNUL (pagg. 92-95)

IPNUL permette la valutazione dell'ipotesi nulla, cioè permette di determinare se una popolazione di dati e un campione di essa siano omogenei o meno: in caso affermativo (ipotesi nulla verificata) si è in presenza di un sistema 'equo'. Con questo programma, applicabile ad una qualsiasi popolazione di dati, l'autore ha potuto dimostrare la non equità del Totocalcio, e quindi l'esistenza di una strategia dello scommettitore (ved. capitolo 1).

I risultati dell'autore possono essere riottenuti utilizzando i dati delle tabelle di pagg. 78-88. Nel nostro caso, fissato un determinato concorso, la 'PROPORZIONE DELLA CARATTERISTICA NELLA POPOLAZIONE P2' è la probabilità della colonna vincente, la 'PROPORZIONE DELLA CARATTERISTICA NEL CAMPIONE P1' è la percentuale di vincitori realmente ottenuti (si ottiene moltiplicando il numero dei vincitori per 100 e dividendo il tutto per il numero delle giocate), mentre il 'NUMERO DI INDIVIDUI NEL CAMPIONE N' è il numero delle colonne giocate.

Utilizzo del programma CONTGZ (pagg. 95-99)

Il programma permette di determinare se due parametri descrittivi di una certa popolazione siano o meno correlati tra loro, o, utilizzando la terminologia del testo, se due variabili siano tra loro dipendenti o indipendenti. La determinazione di ciò è estremamente utile, poichè permette l'immediato rilevamento delle variabili significative per la descrizione di un certo fenomeno.

Creare un modello matematico di un generico ente fisico A (quale ad esempio il Totocalcio) significa determinare, in base a processi matematici, una funzione che correli la variabile obiettivo A ad un determinato numero di variabili causa X, Y, Z,... nel seguente modo:

$$A = f(X, Y, Z, \ldots)$$

CONTGZ permette di determinare l'esistenza di correlazione tra la variabile 'obiettivo' (quale è A) ed una possibile variabile 'causa', in modo da poter stabilire subito le quantità significative (cioè X, Y, Z,...) da prendere in considerazione (ved. capitolo 2). Dopo, sarà compito del progettista stabilire la funzione f che lega tutte le variabili causa.

Il programma è quindi utile soprattutto nella prima fase di elaborazione di un modello, quando ancora non sono note le grandezze che devono essere tenute in considerazione per la descrizione del sistema. L'autore ha voluto inserire CONTGZ per poter fornire ai lettori interessati uno strumento adatto per la creazione di un sistema personale.

Vediamo ora un piccolo esempio di applicabilità del programma. Supponiamo di voler commercializzare un prodotto dietetico e di essere interessati a conoscere il tipo di acquirente cui rivolgersi. Le variabili che possono essere correlate al peso di una persona (che diventa variabile 'obiettivo' A) sono tante, tra cui il sesso, la classe sociale di appartenenza o l'età (possibili variabili 'effetto'). Consideriamo per esempio l'età: la domanda è se esiste una correlazione tra il peso di una persona ed i suoi anni.

Supponiamo quindi di aver effettuato una indagine statistica e di aver ottenuto i seguenti dati:

		ETÀ (in anni)			
		0-20	20-40	40-60	60 +
	fino 70	130	80	40	10
PESO	70-80	80	66	53	40
(in kg)	80-90	40	53	66	80
	90 +	10	40	80	130

Dalla tabella (le cifre si riferiscono al numero di persone intervistate aventi le caratteristiche citate) può non essere chiara l'esistenza di correlazione tra peso ed età, che comunque il programma CONTGZ determina immediatamente.

Dato il 'RUN' al programma, bisogna inserire il 'FATTORE A): NUMERO DI RIGHE NR', che nel nostro caso è 4 (il riferimento è alla variabile 'peso'), ed il 'FATTORE B): NUMERO DI COLONNE NC' che è ancora 4 (bisogna ora considerare l'altra variabile); quindi vanno forniti al calcolatore i dati uno per volta, riga dopo riga. Al termine un opportuno messaggio rivelerà l'esistenza di dipendenza.

Spero che la semplicità e l'ingenuità dell'esempio siano perdonati: non si vuole certo suggerire una rivoluzionaria strategia di commercializzazione!

Va infine fatto notare che CONTGZ non accetta matrici di qualunque grandezza, ma solo quelle per cui il prodotto:

(num. colonne - 1) x (num. righe - 1)

sia al massimo dieci, quindi vanno bene matrici da almeno 2 x 2 rino a 2 x 11, 3 x 6 e 4 x 4 (quella dell'esempio), con tutte le possibilità intermedie.

Utilizzo del programma TOTOBRIER (pagg. 99-101)

Il programma permette di determinare, naturalmente a posteriori, la 'bravura' di un dato pronosticatore, mediante il sistema spiegato nel capitolo 7. TOTOBRIER permette quindi la valutazione del migliore pronosticatore nel caso ce ne siano più di uno, senza però dare una valutazione assoluta: prima di definire 'brocco' un pronosticatore è infatti consigliabile offrirgli molte prove di appello. Lo stesso autore ha utilizzato TOTOBRIER per valutare il pronosticatore della 'Repubblica' durante quasi tutto l'arco del campionato, ottenendo un voto-errore più che soddisfacente.

Utilizzando il programma, alla richiesta 'COLONNA VINCENTE F = FINE:' va scritta o la colonna di una determinata settimana nella forma (es.) 1X211X12XX11 se si vuole determinare il voto-errore (nel programma viene chiamato 'ERRORE MEDIO') attribuito al pronostico, oppure va data la lettera 'F' per ottenere il risultato. Da notare che possono essere fornite più colonne all'interno della stessa esecuzione, aumentando in questo modo l'affidabilità del risultato.

Durante l'inserimento dei pronostici, di ogni colonna vanno fornite, in ordine, le probabilità di uscita dei segni 1, X e 2 espresse in forma normalizzata (cioè le somme delle tre probabilità deve essere uno). Se quindi le probabilità sono (es.) 50%, 35% e 15%, bisogna battere:

```
.5          <RETURN>
.35         <RETURN>
.15         <RETURN>
```

oppure equivalentemente:

```
.5,.35,.15  <RETURN>
```

Il risultato viene fornito segnalando sia il voto-errore (o errore medio) che un opportuno commento chiarificatore. La prima quantità non può essere maggiore di 1.333333: in caso contrario è stato commesso un errore nella fase di inserimento dei dati.

Utilizzo del programma TOTOFIN2 (pagg. 101-112)

Questo è forse il programma più importante, in quanto permette l'utilizzo del sistema proposto dall'autore per la determinazione, in base ad un pronostico, delle colonne più probabili in rapporto alle puntate degli scommettitori.

TOTOFIN2 fornisce le colonne sottopuntate dagli scommettitori, basandosi sulle probabilità di uscita dei singoli segni; il numero delle colonne generate dipende da alcuni parametri forniti dall'utente. Vedremo ora l'utilizzo del programma spiegando in ogni momento cosa sta facendo il computer e quali sono i dati da fornirgli.

Dato il 'RUN' al programma, vanno inserite le probabilità di uscita dei segni 1, X e 2 di ogni colonna. Le probabilità devono essere fornite normalizzate (ved. utilizzo del programma TOTOBRIER). Al termine dell'operazione, che deve essere eseguita con estrema accuratezza, il programma utilizza i valori introdotti per la costruzione della tabella di figura 3 (pag. 23): ciò viene segnalato dalle scritte:

INIZIA ROUTINE MONTECARLO
CIRCA 141 COLONNE

ove 141 sono le colonne che il programma accetta. Contemporaneamente viene mostrato un numero progressivo: esso indica il numero di colonne generate dall'inizio. Di ogni colonna viene calcolata la probabilità: se questa è zero (ciò avviene per i limiti di rappresentazione dei numeri della macchina) viene scartata. Quindi il numero totale di colonne generate sarà alla fine ben maggiore di 141. la routine termina o al riempimento della matrice delle colonne accettate o dopo aver generato 2000 colonne, quindi appare il messaggio:

INIZIA COSTRUZIONE CURVA

Per la descrizione della curva e il suo utilizzo rimandiamo alla trattazione dell'autore nel capitolo 3.

Dopo ciò inizia la serie di richieste; la prima delle quali è preceduta dal messaggio:

INSERITE I SEGNI DESIDERATI PER
CIASCUNA PARTITA

Qui (ved. capitolo 8, pag. 65) bisogna indicare i segni desiderati per ogni partita, in modo da stabilire le colonne dalle quali il programma selezionerà quelle sottopuntate. Come afferma l'autore è necessario determinare un compromesso tra il numero di colonne da valutare e il tempo impiegato, considerando che il computer genera circa 42 colonne al minuto (ved. linee 2991-3001 del programma).

Terminata l'operazione, bisogna definire la:

GRIGLIA DI SELEZIONE

mediante la definizione di tre parametri (ved. capitolo 9). La prima richiesta è:

RIDUZIONE 0 = NO   1 = SI

che permette (digitando la cifra '1') di eliminare le colonne che hanno un unico segno diverso da una già accettata. Il secondo parametro richiesto è:

PROBAB. MINIMA ACCETTATA

o 'probabilità di rovina' come è chiamata nel testo. Con questo parametro è possibile non considerare colonne di probabilità troppo bassa: è vero che il sistema TOTOMAC seleziona solo le colonne meno probabili tra le più probabili (non è un gioco di parole!), ma è opportuno non arrivare a considerare quelle le cui possibilità di uscita siano troppo basse. Il valore consigliato dall'autore è .000001 (o 1E-6), variabile tra .0000005 (5E-7) e .00002 (2E-5).

Infine l'ultimo parametro da definire è il:

RAPPORTO TOTOMAC MINIMO

che nel libro viene chiamato 'valore atteso'.

Con questo parametro possiamo considerare solo quelle colonne che permettono una vincita sostanziosa; il suo valore può variare tra 1.1 e 2.

Da questo momento in poi... non è più lavoro nostro! Il computer segnalerà il:

TOTALE DA GENERARE xx

cioè il numero totale xx di colonne che deve considerare (come visto xx dipende dai segni accettabili forniti precedentemente).
Alla richiesta:

CR =

bisogna semplicemente premere il tasto 'RETURN': il computer è in attesa del segnale di via da parte dell'utente (CR sta per l'inglese 'Carriage Return', che indica proprio la pressione del tasto 'RETURN').

Quindi il computer mostra il numero ternario delle colonne che considera (ved. capitolo 8, pagg. 65-66. Non sempre il numero è un intero poichè il C64 è un calcolatore e non una calcolatrice, e quindi non è molto preciso nel calcolo delle potenze dei numeri: ad esempio 3 elevato a 4 dà 81.0000001!). Nel caso una colonna abbia una probabilità maggiore di quella minima accettabile (la 'probabilità di rovina' prima definita), viene attivato il SOTTOMODULO 2.2 (ved. capitolo 8; nel programma alle linee 2455-2530) che verifica se la colonna va bene o meno, segnalando alla fine la situazione delle colonne valutate.

Al termine vengono stampate in successione tutte le colonne accettate (premere 'RETURN' alla richiesta CR). La sequenza non è da relazionare alla probabilità di uscita, dipendendo solamente dall'ordine con cui sono state considerate le colonne dal programma.

Nel caso si ottenga un'unica colonna di soli '1', significa che sono stati forniti valori troppo riduttivi ai tre parametri della griglia di selezione e l'unica cosa da fare è ripetere tutto dall'inizio.

Invece, nel caso di messaggio di errore BAD SUBSCRIPT, il programma ha accettato più di 200 colonne, massimo possibile, ed è quindi necessario rieseguire il programma fornendo valori più riduttivi ai parametri.

Infine per ottenere su stampante le colonne accettate, è necessario introdurre, prima di dare il 'RUN' al programma le seguenti linee:

```
  21  OPEN 4,4
5421  PRINT#,"COLONNA:";Z$
5610  CLOSE 4
```

le quali possono essere utilizzate con una qualsiasi stampante compatibile Commodore.

Utilizzo del programma TOTOTARA (pag. 112-114)

L'ultimo programma allegato al libro permette la taratura del pronosticatore (ved. capitolo 6), determinando eventuali errori sistematici nella determinazione delle percentuali assegnate alle partite. È cioè possibile arrivare a determinare se il pronosticatore scelto abbia o meno la capacità di discriminare con precisione le singole percentuali. Come avverte l'autore, è però necessario avere a disposizione una grande quantità di dati, molto maggiore di quella che si può ottenere in una singola stazione.

Dati negativi forniti dal computer (es. -1) non sono da considerarsi, in quanto corrispondono a percentuali non utilizzate dal pronosticatore.

Finito di stampare
nel mese di Giugno 1988
presso Grafica 85 - Milano

il sistema TOTOMAC

LA NUOVA FRONTIERA DEL TOTOCALCIO PER C64

Totomac

Accattivarsi e dominare la fortuna è da sempre il sogno di ogni uomo, anche non giocatore d'azzardo. Riuscire a prevedere il futuro è d'altra parte grande aspirazione umana.
Il gioco del Totocalcio stimola a dimostrare tali capacità e Totomac affronta entrambi i compiti con una modalità originale: non inventa magie od esorcismi per accattivarsi la dea bendata, non è interessato a questa donna senza volto. Egli guarda in faccia e sviscera il comportamento dei rivali che gli contendono il bottino, e si inserisce con astuzia là dove "sistemisti" o "trottolisti" peccano di faciloneria e di ingenuità. Uomo del nostro tempo, Totomac si affida all'ausilio del computer per individuare le probabilità meno utilizzate dalla massa dei giocatori, e, avendo ben presente i principi della selezione naturale, cerca di batterli con tecniche intelligenti.
Ma se, in futuro, tutti useranno i suoi metodi, cos'altro bisognerà inventare per battere la massa? Totomac non ce lo dice. Forse sta già pensando a nuove tecniche per battere anche se stesso. E se non ci penserà lui, ci penserà qualcun altro. Perché la voglia di vincere non finisce qui.

- Il modello Totomac
- Il pronostico probabilistico
- Taratura e valutazione
- Il programma in esecuzione
- Controllo dei parametri
- Valore atteso
- Probabilità di rovina
- Riduzione del sistema

GRUPPO EDITORIALE JACKSON

L. 29.000 Cod. 576D ISBN 88-7056-220-4